

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 6.45098 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-9
Perfect score: 120
Sequence: 1 VGNRGGEWTCIAYSQLRDQCI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	749	Q72391	Q72391 homo sapien
2	120	100.0	1103	Q6MZ4	Q6MZ4 homo sapien
3	120	100.0	2193	Q6MZ7	Q6MZ7 homo sapien
4	120	100.0	2240	Q6MDP8	Q6MDP8 homo sapien
5	120	100.0	2267	Q6MDP9	Q6MDP9 homo sapien
6	120	100.0	2296	Q6N0A6	Q6N0A6 homo sapien
7	120	100.0	2357	Q6BDT4	Q6BDT4 homo sapien
8	120	100.0	2386	1 F1NC HUMAN	P02751 homo sapien
9	120	100.0	2444	Q6N025	Q6N025 homo sapien
10	120	100.0	2477	Q6MZU5	Q6MZU5 homo sapien
11	119	99.2	2265	1 F1NC_BOVIN	P07589 bos taurus
12	115	95.8	141	2 Q90XQ2	Q90XQ2 ambystoma m
13	114	95.0	190	1 F1NC_NOTVI	Q91400 notophthalm
14	113	94.2	922	2 Q93405	Q93405 brachydanio
15	113	94.2	2478	2 Q93406	Q93406 brachydanio
16	112	93.3	2481	1 F1NC_XENLA	Q91740 xenopus lae
17	112	93.3	2481	2 Q6GQ45	Q6GQ45 xenopus lae
18	110	91.7	2477	1 F1NC_MOUSE	P11276 mus musculus
19	107	89.2	2477	1 F1NC_RAT	P04937 rattus norv
20	96	80.0	2408	2 Q6JAN2	Q6JAN2 brachydanio
21	57.5	47.9	296	2 Q8C6J7	Q8C6J7 mus musculus
22	53	44.2	768	2 Q8YE51	Q8YE51 brucella me
23	53	44.2	859	2 Q8FX25	Q8FX25 brucella su
24	53	44.2	866	2 Q9FH19	Q9FH19 arabidopsis
25	52	43.3	322	2 Q20990	Q20990 caenorhabdi
26	51.5	42.9	289	1 YM01 CAEEL	P34492 caenorhabdi
27	51	42.5	143	2 Q9CZ77	Q9CZ77 mus musculus
28	50	41.7	140	2 Q99A64	Q99A64 bovine vira
29	50	41.7	140	2 Q99A65	Q99A65 bovine vira
30	50	41.7	140	2 Q99A69	Q99A69 bovine vira
31	50	41.7	140	2 Q99A70	Q99A70 bovine vira

32 49.5 41.2 361 2 Q8KQ71 Q8KQ71 uncultured
33 49.5 41.2 420 2 Q9C6F6 Q9C6F6 arabidopsis
34 49.5 41.2 420 2 Q9C6R5 Q9C6R5 arabidopsis
35 49 40.8 140 2 Q99A21 Q99A21 bovine vira
36 49 40.8 140 2 Q99A37 Q99A37 bovine vira
37 49 40.8 143 2 Q8BDJ2 Q8BDJ2 pseudomonas
38 49 40.8 188 2 Q8XYN7 Q8XYN7 raietonia s
39 49 40.8 556 2 Q8BL21 Q8BL21 mus musculus
40 49 40.8 600 2 Q9EPM6 Q9EPM6 mus musculus
41 49 40.8 712 2 Q8BIJ7 Q8BIJ7 mus musculus
42 49 40.8 844 2 Q9JH67 Q9JH67 ovine astro
43 48 40.0 568 2 Q8PMCO Q8PMCO xanthomonas
44 48 40.0 718 2 Q9H2T6 Q9H2T6 pseudomonas
45 48 40.0 720 2 Q9LDM9 Q9LDM9 nicotiana t

ALIGNMENTS

RESULT 1

Q72391 PRELIMINARY; PRT; 749 AA.
AC Q72391,
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Human colon endothelial primary cell culture;
RA Bloecher H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -.
DR HSSP; Q96KP7; 1PBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDP97F3ED2F0DE CRC64;

Query Match 100.0%; Score 120; DB 2; Length 749;

Best Local Similarity 100.0%; Pred. No. 6.2e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGNRGGEWTCIAYSQLRDQCI 21

Db 591 VGNRGGEWTCIAYSQLRDQCI 611

RESULT 2

Q6MZ4 PRELIMINARY; PRT; 1103 AA.

AC Q6MZ4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219) (Fragment).

DR	InterPro; IPR008957; FN_III-like.
DR	InterPro; IPR000562; FN_Type_II.
DR	Pfam; PF00039; fn1; 7.
DR	Pfam; PF00040; fn2; 2.
DR	Pfam; PF00041; fn3; 17.
DR	PRINTS; PR00012; ENTYPEI.
DR	PRINTS; PR00013; ENTYPEII.
DR	PRINTS; PR00014; ENTYPEIII.
DR	ProDom; PD000995; FN_Type_II; 2.
DR	SMART; SM00058; FN1; 7.
DR	SMART; SM00059; FN2; 2.
DR	SMART; SM00060; FN3; 17.
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR	PROSITE; PS01253; FIBRONECTIN_1; 7.
DR	PROSITE; PS00023; FIBRONECTIN_2; 2.
DR	PROSITE; PS00853; FN3; 17.
DR	Hypothetical protein.
FT	NON_TER
SQ	SEQUENCE 2193 AA; 240641 MW; F876E93106540EF3 CRC64;
 Query Match 100.0%; Score 120; DB 2; Length 2193; Best Local Similarity 100.0%; Pred. No. 1.9e-09; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 VGNRGGEWTCIAYSQLRDQC 21
DB	246 VGNRGGEWTCIAYSQLRDQC 266
 RESULT 4	
Q68DP8	
ID	Q68DP8 PRELIMINARY; PRT; 2240 AA.
AC	Q68DP8;
DT	25-OCT-2004 (TrEMBLrel. 28, Created)
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE	Hypochemical protein DKFZp686H0342.
GN	Name=DKFZp686H0342;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;	
RP	[1]
RN	SEQUENCE FROM N.A.
RC	TISSUE=Amygdala;
RG	The German cDNA Consortium;
RA	Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaipe A.,
RA	Mewes H.W., Weil B., Amid C., Oranger A., Fobo G., Han M., Wiemann S.;
RL	Submitted (AUG-2004) to the EMBL/GenBank/DBD databases.
DR	EMBL; CR749317; CAH18172.1; -
DR	InterPro; IPR002086; Aldehyde dehydrog.
DR	InterPro; IPR006209; EGF like.
DR	InterPro; IPR000083; Fibnctnl.
DR	InterPro; IPR003961; FN_III.
DR	InterPro; IPR008957; FN_III-like.
DR	InterPro; IPR000562; FN_Type_II.
DR	Pfam; PF00039; fn1; 12.
DR	Pfam; PF00040; fn2; 2.
DR	Pfam; PF00041; fn3; 15.
DR	PRINTS; PR00012; ENTYPEI.
DR	PRINTS; PR00013; ENTYPEII.
DR	PRINTS; PR00014; ENTYPEIII.
DR	ProDom; PD000995; FN_Type_II; 2.
DR	SMART; SM00058; FN1; 12.
DR	SMART; SM00059; FN2; 2.
DR	SMART; SM00060; FN3; 15.
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR	PROSITE; PS01253; FIBRONECTIN_1; 10.
DR	PROSITE; PS00023; FIBRONECTIN_2; 2.
KW	Hypothetical protein.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 1.59216 Seconds
(without alignments)
1369.066 Million cell updates/sec

Title: US-09-581-651D-9

Perfect score: 120

Sequence: 1 VGNRGSEWTCIAYSLRDQCI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	99.2	2265	1 FNBO	fibronectin - bovi
2	115	95.8	2386	1 FNHU	fibronectin precu
3	114	95.0	190	2 I51279	fibronectin - east
4	112	93.3	2481	2 A43908	fibronectin - Afri
5	107	89.2	2477	2 S14428	fibronectin precu
6	53	44.2	768	2 A33505	alkaline phosphata
7	52	43.3	143	2 T28845	hypothetical prote
8	51.5	42.9	786	2 S44837	K02D10.1 protein -
9	49.5	41.2	420	2 S86614	hypothetical prote
10	48	40.0	718	2 A83282	probable TonB-depe
11	47.5	39.6	405	2 T12805	hypothetical prote
12	47	39.2	295	2 T49230	hypothetical prote
13	47	39.2	502	1 J02341	arginine decarboxy
14	47	39.2	612	2 T35430	probable long-chai
15	46	38.3	315	2 A64205	O-sialoglycoprotei
16	45.5	37.9	416	2 T10623	hypothetical prote
17	45	37.5	172	2 S16208	NADH2 dehydrogenas
18	45	37.5	355	2 B95213	ABC transporter, s
19	45	37.5	355	2 C98077	hypothetical prote
20	45	37.5	621	2 T15046	arginine decarboxy
21	44.5	37.1	366	2 S11363	actin modulator pr
22	44.5	37.1	780	2 T00366	hypothetical prote
23	44	36.7	147	2 S64252	probable membrane
24	44	36.7	256	2 C70687	hypothetical prote
25	44	36.7	319	2 S73421	o-sialoglycoprotei
26	44	36.7	437	1 FGHUG	fibronectin gamma-A
27	44	36.7	453	1 FGHUGB	fibronectin gamma-B
28	44	36.7	583	2 T25690	hypothetical prote
29	44	36.7	700	2 T27364	hypothetical prote

ALIGNMENTS

RESULT 1

FNBO

fibronectin - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004

C:Accession: A26452; B21165; A23292

R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.

Eur. J. Biochem. 161, 441-453, 1986

A:Title: Complete primary structure of bovine plasma fibronectin.

A:Reference number: A26452; MUID:87054047; PMID:3780752

A:Accession: A26452

A:Molecule type: protein

A:Residues: 1-2265 <SKO>

A:Cross-references: UNIPROT:P07589

R:Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.

Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983

A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins

A:Reference number: A21165; MUID:83221567; PMID:6304699

A:Accession: B21165

A:Molecule type: mRNA

A:Residues: 2170-2265 <KOR>

A:Cross-references: GB:K00800; NID:g163055; PIDN:AAA30521.2; PID:g5713333

R:Petersen, T.E.; Thøgersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sottl

Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983

A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal

A:Reference number: A23292; MUID:83117805; PMID:6218503

A:Accession: A23292

A:Molecule type: protein

A:Residues: 1-16, 'C', 18-20, 'S', 22-432, 447-463, 1367-1517, 1567-1673, 2062-2176, 'N', 2178-226

C:Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.

C:Comment: The plasma fibronectin molecule consists of two chains, which are connected L

C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib

aling, and maintenance of cell shape.

C:Comment: Plasma fibronectin is synthesized by hepatocytes.

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe

C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu

F:21-241/Domain: fibrin and heparin binding <FB>

F:21-56/Domain: fibronectin type I repeat homology <1F1>

F:66-104/Domain: fibronectin type I repeat homology <1F2>

F:110-148/Domain: fibronectin type I repeat homology <1F3>

F:155-194/Domain: fibronectin type I repeat homology <1F4>

F:200-239/Domain: fibronectin type I repeat homology <1F5>

F:277-577/Domain: collagen binding <CB>

F:277-311/Domain: fibronectin type I repeat homology <1F6>

F:329-370/Domain: fibronectin type II repeat homology <2F1>

F:389-430/Domain: fibronectin type II repeat homology <2F2>

F:439-477/Domain: fibronectin type I repeat homology <1F7>

F:487-524/Domain: fibronectin type I repeat homology <1F8>

F:530-568/Domain: fibronectin type I repeat homology <1F9>

F:578-661/Domain: fibronectin type III repeat homology <FN3A>

F:688-770/Domain: fibronectin type III repeat homology <FN3B>

F:779-860/Domain: fibronectin type III repeat homology <FN3C>

genome polypeptide
hypothetical prote
probable membrane
maltose permease -
hypothetical prote
probable phosphogl
hypothetical prote
hypothetical prote
surface glycoprote
sugar permease hom
RNA helicase DbpA
surface glycoprote
genome polypeptide
probable dimethyl
protein F28C11.9
hypothetical prote

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 7.1098 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-9

Perfect score: 120

Sequence: 1 VGNRGWETCIAYSQLRDQCI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	343	9	US-09-934-706-1 Sequence 1, Appli
2	120	100.0	400	9	US-09-934-706-5 Sequence 5, Appli
3	120	100.0	501	9	US-09-934-706-4 Sequence 4, Appli
4	120	100.0	642	16	US-10-741-601-354 Sequence 354, App
5	120	100.0	642	17	US-10-741-600-1066 Sequence 1066, Ap
6	120	100.0	657	16	US-10-741-601-359 Sequence 359, App
7	120	100.0	657	17	US-10-741-600-1072 Sequence 1072, Ap
8	120	100.0	984	16	US-10-741-601-356 Sequence 356, App
9	120	100.0	984	17	US-10-741-600-1069 Sequence 1069, Ap
10	120	100.0	1173	18	US-10-450-763-52634 Sequence 52634, A
11	120	100.0	2220	15	US-10-236-392-4 Sequence 4, Appli

12	120	100.0	2296	16	US-10-741-601-363 Sequence 363, App
13	120	100.0	2296	17	US-10-741-600-1075 Sequence 1075, Ap
14	120	100.0	2320	14	US-10-279-733-8 Sequence 8, Appli
15	120	100.0	2320	15	US-10-236-392-2 Sequence 2, Appli
16	120	100.0	2328	14	US-10-171-311-64 Sequence 64, Appli
17	120	100.0	2328	15	US-10-236-031B-70 Sequence 70, Appli
18	120	100.0	2328	15	US-10-374-979-98 Sequence 98, Appli
19	120	100.0	2328	15	US-10-182-936A-98 Sequence 98, Appli
20	120	100.0	2328	16	US-10-477-238A-677 Sequence 677, App
21	120	100.0	2328	16	US-10-680-287A-677 Sequence 677, App
22	120	100.0	2328	17	US-10-477-173-677 Sequence 104, App
23	120	100.0	2355	15	US-10-144-194A-104 Sequence 235, App
24	120	100.0	2355	15	US-10-360-101-235 Sequence 3, Appli
25	120	100.0	2355	15	US-10-447-161-3 Sequence 94, Appli
26	120	100.0	2355	16	US-10-734-564-94 Sequence 357, App
27	120	100.0	2355	16	US-10-741-601-357 Sequence 357, App
28	120	100.0	2355	16	US-10-741-601-366 Sequence 366, App
29	120	100.0	2355	16	US-10-491-566-104 Sequence 104, App
30	120	100.0	2355	17	US-10-741-600-1067 Sequence 1067, Ap
31	120	100.0	2355	17	US-10-741-600-1078 Sequence 1078, Ap
32	120	100.0	2355	17	US-10-852-335A-147 Sequence 147, App
33	120	100.0	2355	18	US-10-287-436A-436 Sequence 436, App
34	120	100.0	2355	18	US-10-287-436A-1137 Sequence 1137, Ap
35	120	100.0	2355	20	US-11-040-130-28 Sequence 28, Appli
36	120	100.0	2386	10	US-09-961-403-1 Sequence 1, Appli
37	120	100.0	2386	16	US-10-741-601-360 Sequence 360, App
38	120	100.0	2386	16	US-10-788-792-206 Sequence 206, App
39	120	100.0	2386	16	US-10-618-281-32 Sequence 32, Appli
40	120	100.0	2386	17	US-10-741-600-1071 Sequence 1071, Ap
41	120	100.0	2386	17	US-10-868-577A-59 Sequence 59, Appli
42	120	100.0	2386	18	US-10-485-758-4 Sequence 4, Appli
43	120	100.0	2386	18	US-10-485-758-9 Sequence 9, Appli
44	120	100.0	2474	18	US-10-450-763-52638 Sequence 52638, A
45	96	80.0	228	15	US-10-344-634-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-934-706-1
; Sequence 1, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 1
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified Human
; OTHER INFORMATION: Fibronectin Collagen-Binding Domain
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN
; LOCATION: (2)..(341)
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: CONFLICT
; LOCATION: (69)
; NAME/KEY: CONFLICT
; LOCATION: (125)
US-09-934-706-1

Query Match 100.0%; Score 120; DB 9; Length 343;

Best Local Similarity 100.0%; Pred. No. 3.9e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 VNGRGWTCIAYSQLRDQCI 21
   |||||
Db 210 VNGRGWTCIAYSQLRDQCI 230

RESULT 2
US-09-934-706-5
; Sequence 5, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 5
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Hybrid
; OTHER INFORMATION: Polypeptide of Human Fibronectin Collagen-Binding
; OTHER INFORMATION: Domain and Human Epidermal Growth Factor
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN
; LOCATION: (2)..(341)
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: PEPTIDE
; LOCATION: (343)..(347)
; OTHER INFORMATION: /note=" enterokinase recognition sequence"
; NAME/KEY: PEPTIDE
; LOCATION: (348)..(400)
; OTHER INFORMATION: /note="human epidermal growth factor"
US-09-934-706-5

Query Match 100.0%; Score 120; DB 9; Length 400;
Best Local Similarity 100.0%; Pred. No. 4.5e-10; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
   |||||
Db 210 VNGRGWTCIAYSQLRDQCI 230

RESULT 3
US-09-934-706-4
; Sequence 4, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Hybrid
; OTHER INFORMATION: Polypeptide of Human Fibronectin Collagen-Binding
; OTHER INFORMATION: Domain and Human Basic Fibroblast Growth Factor
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN

```

```

; LOCATION: (2)..(341)
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: PEPTIDE
; LOCATION: (343)..(347)
; OTHER INFORMATION: /note="enterokinase recognition sequence"
; NAME/KEY: PEPTIDE
; LOCATION: (348)..(501)
; OTHER INFORMATION: /note="human fibroblast growth factor"
US-09-934-706-4

Query Match 100.0%; Score 120; DB 9; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
   |||||
Db 210 VNGRGWTCIAYSQLRDQCI 230

RESULT 4
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match 100.0%; Score 120; DB 16; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
   |||||
Db 484 VNGRGWTCIAYSQLRDQCI 504

RESULT 5
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match 100.0%; Score 120; DB 17; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
   |||||
Db 484 VNGRGWTCIAYSQLRDQCI 504

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 2.22353 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-9

Perfect score: 120

Sequence: 1 VGNRGWEWCIAYSQLRDQCI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pdp.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pdp.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pdp.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pdp.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pdp.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	2324	1 US-08-283-857-1	Sequence 1, Appli
2	120	100.0	2324	5 PCT-US95-09819-1	Sequence 1, Appli
3	120	100.0	2327	6 5455158-1	Patent No. 5455158
4	120	100.0	2327	6 5455158-1	Patent No. 5455158
5	120	100.0	2386	4 US-09-961-403-1	Sequence 1, Appli
6	120	100.0	2446	2 US-08-551-356-2	Sequence 2, Appli
7	120	100.0	2446	5 PCT-US93-12887-2	Sequence 2, Appli
8	115	95.8	2231	1 US-08-153-799-16	Sequence 16, Appli
9	115	95.8	2386	2 US-08-016-366A-12	Sequence 20, Appli
10	102	85.0	48	2 US-08-982-597A-20	Sequence 12, Appli
11	102	85.0	48	3 US-09-136-218-20	Sequence 20, Appli
12	50	43.7	115	4 US-09-902-540-12185	Sequence 12185, A
13	48	40.0	723	4 US-09-252-991A-18279	Sequence 18279, A
14	46	38.3	231	4 US-09-328-352-6328	Sequence 6328, Ap
15	45	37.5	355	4 US-09-583-110-3026	Sequence 3026, Ap
16	45	37.5	729	4 US-09-543-681A-8257	Sequence 8257, Ap
17	44	36.7	20	4 US-09-664-945-4	Sequence 4, Appli
18	44	36.7	27	4 US-09-664-945-65	Sequence 65, Appli
19	44	36.7	122	4 US-09-513-999C-7822	Sequence 7822, Ap
20	44	36.7	147	4 US-09-538-092-334	Sequence 324, App
21	44	36.7	206	4 US-09-902-540-16467	Sequence 16467, A
22	44	36.7	225	4 US-09-270-767-60597	Sequence 60597, A
23	44	36.7	302	4 US-09-270-767-45103	Sequence 45103, A
24	44	36.7	449	4 US-09-949-016-9614	Sequence 9614, Ap
25	44	36.7	449	4 US-09-949-016-9615	Sequence 9615, Ap
26	44	36.7	453	1 US-08-206-176-6	Sequence 6, Appli
27	44	36.7	454	3 US-08-434-099A-27	Sequence 27, Appli

Sequence 28, Appli
Sequence 8593, Ap
Sequence 8594, Ap
Sequence 9447, Ap
Sequence 6880, Ap
Sequence 16119, A
Sequence 27, Appli
Sequence 4577, Ap
Sequence 37136, A
Sequence 52353, A
Sequence 821, App
Sequence 3406, Ap
Sequence 2, Appli
Sequence 18154, A
Sequence 7905, Ap
Sequence 21, Appli
Sequence 21, Appli

ALIGNMENTS

RESULT 1
US-08-283-857-1
; Sequence 1, Application US/08283857
; Patent No. 5792742
; GENERAL INFORMATION:
; APPLICANT: GOLD, Leslie I.
; APPLICANT: ROSTAGNO, Agueda A.
; APPLICANT: BARON, Martin
; APPLICANT: CAMPBELL, Iain D.
; APPLICANT: WILLIAMS, Michael, J.
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USERS THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/283.857
; FILING DATE: 01-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/714,134
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-283-857-1

Query Match 100.0%; Score 120; DB 1; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2e-09;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
|||||
Db 468 VNGRGWTCIAYSQLRDQCI 488

RESULT 2

PCT-US95-09819-1
; Sequence 1, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-09819-1

Query Match 100.0%; Score 120; DB 5; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
|||||
Db 468 VNGRGWTCIAYSQLRDQCI 488

RESULT 3

5455158-1
; Patent No. 5455158
; APPLICANT: VOGEL, TIKVA; LEVANON, AVIGDOR; WERBER, MOSHE M.;
; GUY, RACHEL; PANET, AMOS
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,241
; FILING DATE: 04-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 526,397
; FILING DATE: 21-MAY-1990

; APPLICATION NUMBER: 345,952
; FILING DATE: 28-APR-1989
; APPLICATION NUMBER: 291,951
; FILING DATE: 29-DEC-1988
; SEQ ID NO:1:
; LENGTH: 2327
5455158-1

Query Match 100.0%; Score 120; DB 6; Length 2327;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
|||||
Db 471 VNGRGWTCIAYSQLRDQCI 491

RESULT 4

5455158-1
; Patent No. 5455158
; APPLICANT: VOGEL, TIKVA; LEVANON, AVIGDOR; WERBER, MOSHE M.;
; GUY, RACHEL; PANET, AMOS
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,241
; FILING DATE: 04-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 526,397
; FILING DATE: 21-MAY-1990
; APPLICATION NUMBER: 345,952
; FILING DATE: 28-APR-1989
; APPLICATION NUMBER: 291,951
; FILING DATE: 29-DEC-1988
; SEQ ID NO:1:
; LENGTH: 2327
5455158-1

Query Match 100.0%; Score 120; DB 6; Length 2327;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
|||||
Db 471 VNGRGWTCIAYSQLRDQCI 491

RESULT 5

US-09-961-403-1
; Sequence 1, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-1

Query Match 100.0%; Score 120; DB 4; Length 2386;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 7.76863 Seconds
(without alignment)
1045.483 Million cell updates/sec

Title: US-09-581-651D-9
Perfect score: 120
Sequence: 1 VGNRGWETCTAYSLRDQCI 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 396760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	100.0	21	2 AAY28906	Aay28906 MSF 1-alp
2	120	100.0	343	3 AAB08505	Aab08505 Amino aci
3	120	100.0	343	5 ABB07961	Abb07961 Modified
4	120	100.0	400	3 AAB08509	Aab08509 Hybrid of
5	120	100.0	461	5 ABB07964	Abb07964 Human fib
6	120	100.0	473	5 ABB07965	Abb07965 Human fib
7	120	100.0	501	3 AAB08508	Aab08508 Hybrid of
8	120	100.0	642	8 ADQ39403	Adq39403 Human myo
9	120	100.0	642	8 ADR67316	Adr67316 Human bla
10	120	100.0	642	8 ADS17489	Ads17489 Amino aci
11	120	100.0	642	8 ADR97658	Adr97658 Human fib
12	120	100.0	657	8 ADQ39409	Adq39409 Human myo
13	120	100.0	660	2 AAY28901	Aay28901 Human mig
14	120	100.0	984	8 ADQ39406	Adq39406 Human myo
15	120	100.0	1173	4 ABG22275	Abg22275 Novel hum
16	120	100.0	1179	8 ADP75952	Adp75952 Human min
17	120	100.0	1223	8 ADP75957	Adp75957 Human sec
18	120	100.0	1359	8 ADP75957	Adp75957 Human leu
19	120	100.0	2182	8 ADR66462	Adr66462 Human pro
20	120	100.0	2182	8 ADR66120	Adr66120 Human pro
21	120	100.0	2220	6 ABO01289	Ab001289 Human pro
22	120	100.0	2220	8 ADN95950	Adn95950 Human NOV
23	120	100.0	2265	4 AAM38647	Aam38647 Human pol
24	120	100.0	2266	6 ABR40124	Abr40124 Human cel
25	120	100.0	2296	8 ADQ39412	Adq39412 Human myo

ALIGNMENTS

RESULT 1

AAY28906
ID AAY28906 standard; peptide; 21 AA.

XX AC AAY28906;

DT 21-SEP-1999 (first entry)

DE MSF 1-alpha peptide epitope.

XX KW Migration stimulatory factor; MSF; cell-migration; modulation; human;
wound healing; scarring; MSF1-alpha; epitope; fibronectin.

XX OS Synthetic.

XX OS Homo sapiens.

XX FN WO9931233-Al.

XX PD 24-JUN-1999.

XX PF 15-DEC-1998; 98WO-GB003766.

XX PR 16-DEC-1997; 97GB-00026539.

XX PA (UYDU-) UNIV DUNDEE.

XX PI Schor SL, Schor AM;

XX DR WPI, 1999-430039/36.

XX PT Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX PS Claim 18; Page 57; 86pp; English.

XX CC The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. Sequences AAY28902-906 represent peptide epitopes of MSF against which monoclonal antibodies that are specific to MSF and do not cross-react with fibronectin are raised

XX SQ Sequence 21 AA;

Query Match 100.0%; Score 120; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;

Aag37107 Human fib
Abr42588 Human fib
Abo01288 Human pro
Adr90519 Human ful
Aar92778 Human fib
Aau74674 Human fib
Aae23651 Human pro
Aap70373 Human fib
Aar15468 Human fib
Aag68182 Fibronect
Adu07486 Protein d
Abr41106 Human fib
Abr92078 Human cer
Adb70378 Fibronect
Adb98726 Human fib
Ade82522 Human pro
Adj37157 Human mal
Aam38646 Human pol
Aam38649 Human pol
Abr58335 NM_00202

26 120 100.0 2320 6 AAE37107
27 120 100.0 2320 6 ABR42588
28 120 100.0 2320 6 ABO01288
29 120 100.0 2320 8 ADR90519
30 120 100.0 2324 2 AAR92778
31 120 100.0 2324 5 AAU74674
32 120 100.0 2324 5 AAE23651
33 120 100.0 2327 1 AAP70373
34 120 100.0 2327 2 AAR15468
35 120 100.0 2328 4 AAG68182
36 120 100.0 2328 6 ABU07486
37 120 100.0 2328 6 ABR41106
38 120 100.0 2328 6 ABR92078
39 120 100.0 2328 7 ADB70378
40 120 100.0 2328 7 ADB98726
41 120 100.0 2328 7 ADE82522
42 120 100.0 2328 8 ADJ37157
43 120 100.0 2330 4 AAM38646
44 120 100.0 2355 4 AAM38649
45 120 100.0 2355 6 ABR58335

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
 DB 1 VNGRGWTCIAYSQLRDQCI 21

RESULT 2
 AAB08505
 ID AAB08505 standard; protein; 343 AA.
 AC AAB08505;
 XX
 XX 20-DEC-2000 (first entry)
 DT
 DE Amino acid sequence of modified fibronectin collagen-binding domain.
 XX
 KW Fibronectin; collagen-binding domain; sustained release; gene therapy;
 KW physiologically active polypeptide; topical retention;
 KW tissue regeneration.
 XX
 XX Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 1 /note= "Met added for efficient expression"
 FT
 XX WO200049159-A1.
 PN
 XX 24-AUG-2000.
 PD
 XX 21-FEB-2000; 2000WO-JP000964.
 PF
 XX 19-FEB-1999; 99JP-00041913.
 PR
 PR 01-NOV-1999; 99JP-00311364.
 XX
 XX (TERU) TERUMO CORP.
 PA
 XX Ishikawa T, Kitajima T;
 PI
 XX WPI; 2000-565375/52.
 DR N-PSDB; AAA64263.
 XX
 XX Collagen-binding active polypeptide for use in an agent for enabling
 PT topical retention or sustained release of a physiologically active
 PT peptide or physiological activity-imparting agent comprises a fibronectin
 PT peptide.
 XX
 XX Disclosure; Page 113-115; 135pp; English.

The present sequence represents a modified fibronectin collagen-binding domain. The protein is used to construct a collagen-binding CC physiologically active polypeptide. This polypeptide comprises a peptide CC from fibronectin ligated to a physiologically active peptide. The CC polypeptides are used in an agent for enabling topical retention or CC sustained release of a physiologically active peptide or physiological CC activity-imparting agent. They may be used in gene therapy and in tissue CC regeneration

XX
 XX Sequence 343 AA;

Query Match 100.0%; Score 120; DB 3; Length 343;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
 DB 210 VNGRGWTCIAYSQLRDQCI 230

RESULT 3
 AAB07961

ID AAB07961 standard; protein; 343 AA.
 XX
 AC AAB07961;
 XX
 XX 12-AUG-2002 (first entry)
 DT
 DE Modified human fibronectin collagen-binding domain.
 XX
 KW Osteogenesis; drug delivery system; DDS; collagen; osteopathic; human;
 KW fibronectin.
 XX
 XX Homo sapiens.
 OS Synthetic.
 XX
 XX JP2002058485-A.
 PN
 XX 26-FEB-2002.
 PD
 XX 16-AUG-2000; 2000JP-00246744.
 PF
 XX 16-AUG-2000; 2000JP-00246744.
 PR
 XX (TERU) TERUMO CORP.
 PA
 XX WPI; 2002-438612/47.
 DR N-PSDB; ABL41020.
 XX
 XX Novel osteogenesis stimulating fused protein having collagen avidity used
 PT as an osteogenesis stimulator and a localizing agent.
 PT
 XX Disclosure; Page 17-18; 30pp; Japanese.

The invention provides an osteogenesis stimulating fused protein (I) for a drug delivery system (DDS) of osteogenetic factor, having collagen CC avidity and polypeptides homologous to collagen avidity domain or its CC modified peptides. (I) is used for stimulation of osteogenesis, a CC localizing agent and a slow releasing agent for a drug delivery system. CC The present sequence represents a modified human fibronectin collagen-binding domain

XX
 XX Sequence 343 AA;

Query Match 100.0%; Score 120; DB 5; Length 343;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
 DB 210 VNGRGWTCIAYSQLRDQCI 230

RESULT 4
 AAB08509
 ID AAB08509 standard; protein; 400 AA.
 XX
 AC AAB08509;
 XX
 XX 12-SEP-2003 (revised)
 DT 20-DEC-2000 (first entry)
 DE Hybrid of fibronectin collagen-binding domain/epidermal growth factor.
 XX
 KW Fibronectin; collagen-binding domain; sustained release; gene therapy;
 KW physiologically active polypeptide; topical retention; chimera;
 KW tissue regeneration; epidermal growth factor.

XX
 XX Homo sapiens.
 OS Homo sapiens.
 OS Chimeric.
 XX
 XX Key Location/Qualifiers
 FH Protein 2..341 /note= "human fibronectin collagen-binding domain"
 FT

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 436.745 Seconds
(without alignments)
1830.243 Million cell updates/sec

Title: US-09-581-651D-9

Perfect score: 120

Sequence: 1 VGNRGWTCIAYSLRDQCI 21

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum Match 100%

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US9581651/runat_07112005_092223_28789/app_query.fasta_1.1834
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CGN 1 1 8180 @runat_07112005_092223_28789 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DE TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:

1: gb_est1.*

2: gb_est2.*

3: gb_hic.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gss1.*

9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	207	1	AL710586 DKFZp686N
c	2	120	100.0	242	5 BQ340610 PMO-NN025
	3	120	100.0	304	1 AA852107 NHTBCae10
	4	120	100.0	321	4 B1036180 CM3-NT026
c	5	120	100.0	369	5 BQ340603 PMO-NN025
	6	120	100.0	373	2 B5817152 PM3-NN025
c	7	120	100.0	380	5 BQ340568 PMO-NN025
c	8	120	100.0	385	5 BQ340613 PMO-NN025
	9	120	100.0	386	7 CN419560 170005321

10	120	100.0	407	7	CN419462
11	120	100.0	436	2	BF868748
12	120	100.0	461	7	CF128990
13	120	100.0	480	7	CN419466
c	14	120	100.0	497	4 EG000888
	15	120	100.0	503	4 CN419452
c	16	120	100.0	525	2 AW853981
	17	120	100.0	530	1 AU141012
	18	120	100.0	534	5 BX642843
c	19	120	100.0	536	4 BG000869
c	20	120	100.0	542	5 BQ340573
	21	120	100.0	592	7 CN419590
	22	120	100.0	613	6 CD613819
	23	120	100.0	618	7 CN481692
	24	120	100.0	618	7 CN482377
	25	120	100.0	632	7 CN482351
	26	120	100.0	635	5 BQ312550
	27	120	100.0	645	7 CN419499
	28	120	100.0	653	7 CN419474
	29	120	100.0	655	7 CN419546
	30	120	100.0	657	7 CN419503
	31	120	100.0	667	1 AU140910
c	32	120	100.0	670	5 BU620723
	33	120	100.0	673	1 AU140450
	34	120	100.0	684	1 AU134619
	35	120	100.0	695	7 CN419477
	36	120	100.0	699	1 AU140889
	37	120	100.0	699	7 CN419497
	38	120	100.0	702	7 CN419472
	39	120	100.0	705	1 AU140899
	40	120	100.0	707	1 AU140730
	41	120	100.0	723	1 CN419622
	42	120	100.0	730	1 AU140834
	43	120	100.0	731	1 AU140526
	44	120	100.0	732	1 AU140747
45	120	100.0	734	6	CD613786

ALIGNMENTS

RESULT 1
AL710586
LOCUS
DEFINITION
ACCSSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL710586
DKFZp686N066 r1 686 (synonym: hicc3) Homo sapiens linear EST 04-SEP-2003
DKFZp686N066 5', mRNA sequence.

AL710586
GI:19693941

EST.

AL710586.1 Homo sapiens (human)

EST.

Unpublished (1999)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by EMBL (European Molecular Biology Laboratories,

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No sl sequence available.

This clone (DKFZp686N066) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

1..207

/organism="Homo sapiens"

/mol_type="mRNA"

```
/db_xref="taxon:9606"
/clone="DKFZp686N066"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hicc3)"
/notes="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"
```

ORIGIN

Alignment Scores:
 Pred. No.: 3 98e-08 Length: 207
 Score: 120.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-09-581-651D-9 (1-21) x AL710586 (1-207)

QY 1 ValGlyAenGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
 |||||
 DB 54 GTTGGGAATGCTGCTGGGAATGGATGTCATTGCTACTCGACGCTTCGAGATCAGTGC 113
 |||||
 QY 21 Ile 21
 |||||
 DB 114 ATT 116

RESULT 2

BQ340610/C BQ340610 242 bp mRNA linear EST 20-MAY-2002
 LOCUS PMO-NN0258-220501-001-e02 NN0258 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 ACCESSION BQ340610
 VERSION BQ340610.1 GI:21001672
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 242)
 AUTHORS Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL

MEDLINE Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 PUBMED 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PMO&t2=PMO-NN0258-
 220501-001-e02&t3=2001-05-22&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 16
 High quality sequence stop: 241.

FEATURES

Location/Qualifiers
 1..242
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="NN0258"
 /note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 4.75e-08 Length: 242
 Score: 120.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0

US-09-581-651D-9 (1-21) x BQ340610 (1-242)

QY 1 ValGlyAenGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
 |||||
 DB 155 GTTGGGAATGCTGCTGGGAATGGATGTCATTGCTACTCGACGCTTCGAGATCAGTGC 96
 |||||
 QY 21 Ile 21
 |||||
 DB 95 ATT 93

RESULT 3

AA852107 304 bp mRNA linear EST 20-JUN-2002
 LOCUS NHTBCae10c08r1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
 DEFINITION
 ACCESSION AA852107
 VERSION AA852107.1 GI:2940700
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 304)
 AUTHORS Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
 Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
 Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
 SGAP: The Skeletal Genome Anatomy Project
 Unpublished (1997)
 Contact: Libin Jia
 Medical Genetics Branch
 National Human Genome Research Institute
 10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1367, USA
 Tel: 301-402-4877
 Fax: 301-496-7157
 Email: libin@helix.nih.gov

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..304

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NHTBCae10c08"

/sex="Female"

/tissue_type="Bone"

/cell_type="Trabecular Bone Cells"

/lab_host="SURE"

/clone_lib="Normal Human Trabecular Bone Cells"

/note="Organ: Hip; Vector: pBluescript; Site 1: EcoRI;

Library constructed by Dr. Marian Young and Dr. Pamela

Gehron Robey (NIDCR)"

ORIGIN

Alignment Scores:
 Pred. No.: 6.15e-08 Length: 304
 Score: 120.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 107.827 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-9
Perfect score: 120
Sequence: 1 VGNRGWTCIAYSLRDQCI 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 413490567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1834
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -GOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09581651@cgn1_1.1041@runat_07112005_092226_28867
-NCPUI=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:**
1: /cgn2_1/pubdata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/pubdata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/pubdata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/pubdata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/pubdata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/pubdata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/pubdata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/pubdata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/pubdata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/pubdata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/pubdata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/pubdata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/pubdata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/pubdata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/pubdata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/pubdata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/pubdata/1/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/pubdata/1/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/pubdata/1/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/pubdata/1/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/pubdata/1/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/pubdata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/pubdata/1/pubpna/US10_NEW_PUB.seq.*
24: /cgn2_6/pubdata/1/pubpna/US10_NEW_PUB.seq.*
25: /cgn2_6/pubdata/1/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/pubdata/1/pubpna/US11_NEW_PUB.seq.*
27: /cgn2_6/pubdata/1/pubpna/US60_NEW_PUB.seq.*
28: /cgn2_6/pubdata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	100.0	358	18	US-10-242-535A-30516
2	120	100.0	358	19	US-10-085-783A-30516
3	120	100.0	1053	9	US-09-934-706-8
4	120	100.0	1189	15	US-10-198-846-11556
5	120	100.0	1224	9	US-09-934-706-16
6	120	100.0	1527	9	US-09-934-706-15
7	120	100.0	2127	17	US-10-210-120-49
8	120	100.0	2127	22	US-10-956-157-4288
9	120	100.0	2127	24	US-10-909-035-49
10	120	100.0	2443	20	US-10-741-601-70
11	120	100.0	2443	22	US-10-741-600-238
12	120	100.0	2488	20	US-10-741-601-75
13	120	100.0	2488	22	US-10-741-600-244
14	120	100.0	3522	24	US-10-450-763-22266
15	120	100.0	4295	18	US-10-144-194A-51
16	120	100.0	4295	21	US-10-491-566-51
17	120	100.0	6510	20	US-10-741-601-72
18	120	100.0	6510	22	US-10-741-600-241
19	120	100.0	6988	19	US-10-236-392-1
20	120	100.0	7361	19	US-10-236-392-3
21	120	100.0	7677	22	US-10-956-157-4995
22	120	100.0	7679	22	US-10-831-704-38
23	120	100.0	7680	9	US-09-964-824A-574
24	120	100.0	7680	15	US-10-171-311-63
25	120	100.0	7680	18	US-10-236-031B-69
26	120	100.0	7680	18	US-10-374-979-75
27	120	100.0	7680	19	US-10-182-936A-75
28	120	100.0	7680	19	US-10-641-643-1289
29	120	100.0	7680	20	US-10-717-597-222
30	120	100.0	7680	20	US-10-788-792-79
31	120	100.0	7680	21	US-10-477-238A-654
32	120	100.0	7680	21	US-10-580-287A-654
33	120	100.0	7680	22	US-10-278-698-88
34	120	100.0	7680	22	US-10-278-698-603
35	120	100.0	7680	22	US-10-843-641A-5877
36	120	100.0	7680	22	US-10-477-173-654
37	120	100.0	7680	22	US-10-852-335A-52
38	120	100.0	7705	18	US-10-447-161-4
39	120	100.0	7705	16	US-10-084-817-2
40	120	100.0	7823	20	US-10-741-601-77
41	120	100.0	7823	22	US-10-741-600-245
42	120	100.0	7848	20	US-10-741-601-78
43	120	100.0	7848	22	US-10-741-600-246
44	120	100.0	7867	14	US-10-098-841-6
45	120	100.0	7935	20	US-10-741-601-74

ALIGNMENTS

RESULT 1
US-10-242-535A-30516
; Sequence 30516, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242.535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340

```

; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30516
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-30516

```

Alignment Scores:		
Pred. No.:	1.49e-10	358
Score:	120.00	21
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0
Query Match:	100.00%	0
DB:	18	0
		0
Length:		358
Matches:		21
Conservative:		0
Mismatches:		0
Indels:		0
Gaps:		0

US-09-581-651D-9 (1-21) x US-10-242-535A-30516 (1-358)

Qy 1 ValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnIleuArgAspGlnCys 20
 |||||
 Db 199 GTTGGGAATGGTCTGTGGGGAATGGACATGTCCTACTCGCAGCTTCGAGATCACTGC 258
 |||||
 Qy 21 Ile 21
 |||||
 Db 259 ATT 261

RESULT 2

```

US-10-085-783A-30516
; Sequence 30516, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods for
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30516
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-30516

```

Alignment Scores:	
Pred. No.:	1,498-10
Score:	120.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	19
Length:	358
Matches:	21
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-581-651D-9 (1-21) x US-10-085-783A-30516 (1-358)

Qy	1	ValGlyAsnGlyArgGlyCysIuTrpThrCysIlealaTyrSerGlnLeuArgAspGlnCys	20
Db	199	GTTGGGAAATGGTGGTGGGGATGGACATGCATTGCCTACTCGAGCTTCGAGATCACTGC	258
Qy	21	Ile 21	
Db	259	ATT 261	

RESULT 3

```

US-09-934-706-8
; Sequence 8, Application US/09934706
; Patent No., US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16

```

Alignment Scores:

Pred. No.:	5.06e-10	Length:	1053
Score:	120.00	Matches:	21
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-581-651D-9 (1-21) x US-09-934-706-8 (1-1053)

[illegible]

RESULT 4

```

US-10-198-846-11556/c
; Sequence 11556, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11556
; LENGTH: 1189
; TYPE: DNA
;

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: November 11, 2005, 08:49:16 ; Search time 21.1373 Seconds
(without alignments)
1625.652 Million cell updates/sec

Title: US-09-581-651D-9
Perfect score: 120
Sequence: 1 VGNRGWETCIAYSLRDQCI 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092224_28801/app_query.fasta.1.1834
-DB=Issued_Patents_NA -QMT=fastcap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCD=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFTW=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651@cgn2_1.1.187@runat_07112005_092224_28801 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	7679	4	US-09-220-132-38
2	120	100.0	7680	4	Sequence 38, Appl
3	120	100.0	7680	5	Sequence 1289, Ap
4	120	100.0	7705	1	PCT-US95-09819-6
5	120	100.0	7705	2	Sequence 6, Appl
6	120	100.0	7705	2	Sequence 16, Appl
7	120	100.0	7705	6	Sequence 16, Appl
8	120	100.0	7705	6	Sequence 16, Appl
9	120	100.0	7803	2	Patent No. 5455158
10	120	100.0	7803	5	Patent No. 5455158
11	107	89.2	986	1	Sequence 1, Appl
12	107	89.2	986	1	Sequence 1, Appl

c 13	57	47.5	69	1	US-08-259-569-30	Sequence 30, Appl
c 14	57	47.5	69	2	US-08-826-885-30	Sequence 30, Appl
c 15	57	47.5	72	1	US-08-259-569-29	Sequence 29, Appl
c 16	57	47.5	72	2	US-08-826-885-29	Sequence 29, Appl
c 17	57	47.5	72	6	5455158-9	Patent No. 5455158
c 18	57	47.5	72	6	5455158-9	Patent No. 5455158
c 19	56	46.7	601	4	US-09-949-016-160046	Sequence 160046,
c 20	56	46.7	68035	4	US-09-949-016-16219	Sequence 16219, A
c 21	55	45.8	5788	4	US-09-949-016-12498	Sequence 12498, A
c 22	55	45.8	5788	4	US-09-949-016-14458	Sequence 14458, A
c 23	55	45.8	5870	4	US-09-949-016-15247	Sequence 15247, A
c 24	54	45.0	601	4	US-09-949-016-160047	Sequence 160047, A
c 25	51.5	42.9	51927	4	US-09-949-016-17347	Sequence 17347, A
c 26	51.5	42.9	51927	4	US-09-949-016-17348	Sequence 17348, A
c 27	51.5	42.9	73853	4	US-09-949-016-12029	Sequence 12029, A
c 28	51	42.5	483	1	US-08-644-664B-8	Sequence 8, Appl
c 29	51	42.5	483	2	US-08-761-277A-8	Sequence 8, Appl
c 30	51	42.5	483	2	US-08-715-808-15	Sequence 15, Appl
c 31	51	42.5	555	6	5225348-7	Patent No. 5225348
c 32	51	42.5	555	6	5225348-7	Patent No. 5225348
c 33	51	42.5	583	6	5225348-8	Patent No. 5225348
c 34	51	42.5	583	6	5225348-8	Patent No. 5225348
c 35	51	42.5	599	6	5225348-9	Patent No. 5225348
c 36	51	42.5	599	6	5225348-9	Patent No. 5225348
c 37	51	42.5	1451	1	US-08-644-664B-11	Sequence 11, Appl
c 38	51	42.5	1451	2	US-08-761-277A-11	Sequence 11, Appl
c 39	51	42.5	1451	3	US-08-848-760B-14	Sequence 14, Appl
c 40	51	42.5	1451	4	US-09-826-025-14	Sequence 14, Appl
c 41	51	42.5	1561	1	US-07-757-536B-1	Sequence 1, Appl
c 42	51	42.5	1561	6	5225348-4	Patent No. 5225348
c 43	51	42.5	1561	6	5225348-4	Patent No. 5225348
c 44	51	42.5	4695	3	US-09-309-572-9	Sequence 9, Appl
c 45	51	42.5	4695	4	US-09-718-096-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1

US-09-220-132-38
; Sequence 38, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-38

Alignment Scores:
Pred. No.: 6.96e-09
Score: 120.00
Matches: 21
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
Query Match: 100.00%
Indels: 0
DB: 4
Gaps: 0

US-09-581-651D-9 (1-21) x US-09-220-132-38 (1-7679)

Qy 1 ValGlyAsnGlyArgGly-GluTrrThrCysAlaLeuTyrSerGlnLeuArgAspGlnCys 20

Db 1415 GTTGGGAATGCTCGTGGGAATGGACATGTCATTGCGCAACTTCGAGATCGAGTC 1474

Qy 21 Ile 21
Db 1475 ATT 1477

RESULT 2
US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: g31396
US-09-023-655-1289

Alignment Scores:
Pred. No.: 6.96e-09 Length: 7680
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-581-651D-9 (1-21) x US-09-023-655-1289 (1-7680)

Qy 1 ValGlyAanGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
Db 1415 GTTGGGAATGTCGTGGGGAATGGACATGCTACTGCTCAACTTCGAGATCAGTGC 1474

Qy 21 Ile 21
Db 1475 ATT 1477

RESULT 3

PCT-US95-09819-6
; Sequence 6, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US95-09819-6

Alignment Scores:
Pred. No.: 6.96e-09 Length: 7680
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-581-651D-9 (1-21) x PCT-US95-09819-6 (1-7680)

Qy 1 ValGlyAanGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
Db 1415 GTTGGGAATGTCGTGGGGAATGGACATGCTACTGCTCAACTTCGAGATCAGTGC 1474

Qy 21 Ile 21
Db 1475 ATT 1477

RESULT 4
US-08-259-569-16
; Sequence 16, Application US/08259569
; Patent No. 5679320
; GENERAL INFORMATION:
; APPLICANT: Vogel, Tikva
; APPLICANT: Levanon, Avigdor
; APPLICANT: Werber, Moshe
; APPLICANT: Guy, Rachel
; APPLICANT: Panet, Amos
; APPLICANT: Hartman, Jacob
; APPLICANT: Shaked, Hadassa
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:30:56 ; Search time 68.8471 Seconds
(without alignments)
1805.663 Million cell updates/sec

Title: US-09-581-651D-9

Perfect score: 120

Sequence: 1 VGNRGSEWTCIAYSLRDQCI 21

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_SPOOL_h/US09581651/runat_07112005_092222_28772/app_query.fasta.1.1834
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=text -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651_QCGN 1 1052 @runat_07112005_092222_28772 -NCPUs=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	120	100.0	1051	ADS19042	Ads19042 Human fib
2	120	100.0	1053	AAA64263	Aaa64263 cDNA enco
3	120	100.0	1053	ABL41020	Abi41020 Modified
4	120	100.0	1189	ACN90406	ACN90406 Breast ca
5	120	100.0	1224	AAA64271	Aaa64271 Fibronect

6	120	100.0	1407	6	ABL41027	Abi41027 Human fib
7	120	100.0	1443	6	ABL41028	Abi41028 Human fib
8	120	100.0	1527	3	AAA64270	Aaa64270 Fibronect
9	120	100.0	1929	13	ADSI17488	Adsi17488 Nucleotid
10	120	100.0	1929	13	ADR97657	Adr97657 Human fib
11	120	100.0	2127	10	ADDI18477	Adi18477 Human pro
12	120	100.0	2147	2	AAAX81299	Aax81299 Human mlg
13	120	100.0	2358	13	ADR67201	Adr67201 Human bla
14	120	100.0	2443	13	ADQ38575	Adq38575 Human SNP
15	120	100.0	2488	13	ADQ38581	Adq38581 Human SNP
16	120	100.0	3522	5	AA86462	Aaa86462 DNA enco
17	120	100.0	3540	12	ADP75954	Adp75954 Human min
18	120	100.0	3669	12	ADP29134	Adp29134 Human sec
19	120	100.0	4080	12	ADP75958	Adp75958 Human leu
20	120	100.0	4295	8	ACCT2037	Acc22037 BC00770 g
21	120	100.0	4860	3	AAA35009	Aaa35009 Human ade
22	120	100.0	6510	13	ADQ38578	Adq38578 Human SNP
23	120	100.0	6816	10	ABX74443	Abx74443 Human cDN
24	120	100.0	6988	9	ACD06169	Acc06169 Human cDN
25	120	100.0	7020	13	ADR90518	Adr90518 Human cDN
26	120	100.0	7242	13	ADR66637	Adr66637 Human pro
27	120	100.0	7242	13	ADR65953	Adr65953 Human pro
28	120	100.0	7361	9	ACD06170	Acc06170 Human cDN
29	120	100.0	7361	12	ADN95949	Adn95949 Human NOV
30	120	100.0	7550	8	ACC00412	Acc00412 Human cel
31	120	100.0	7677	13	ACN37820	Acn37820 Tumour-as
32	120	100.0	7679	10	ADB31322	Adb31322 Testoster
33	120	100.0	7680	2	AAT17551	Aat17551 Human fib
34	120	100.0	7680	3	AAF21131	Aaf21131 Human low
35	120	100.0	7680	5	ABA82689	Aba82689 Fibronect
36	120	100.0	7680	6	ABL67540	Abi67540 Thyroid c
37	120	100.0	7680	6	ABT11082	Abt11082 Human bre
38	120	100.0	7680	8	ACF03878	Acc03878 Human fib
39	120	100.0	7680	8	ABX10391	Abx10391 DNA enco
40	120	100.0	7680	8	ACC46009	Acc46009 Human fib
41	120	100.0	7680	8	ACF12859	Acc12859 Human gar
42	120	100.0	7680	8	ACA64817	Acc64817 Human fib
43	120	100.0	7680	8	ACA64819	Acc64819 Human fib
44	120	100.0	7680	9	ADB70377	Adb70377 Fibronect
45	120	100.0	7680	10	ADB98703	Adb98703 Human fib

ALIGNMENTS

RESULT 1
ADS19042
ID ADS19042 standard; DNA; 1051 BP.

XX AC ADS19042;

XX DT 16-DEC-2004 (first entry)

XX DE Human fibronectin collagen binding domain DNA SeqID 3.

XX KW human; ds; hepatocyte growth factor; HGF; collagen binding domain;
XX KW fibronectin; artificial blood vessel; stent; tissue regeneration;
XX KW vulnary.

XX OS Homo sapiens.

XX PN JP2004269423-A.

XX PD 30-SEP-2004.

XX PF 07-MAR-2003; 2003JP-00062169.

XX PR 07-MAR-2003; 2003JP-00062169.

XX PA (TERU) TERUMO CORP.

XX DR WPI; 2004-682709/67.

XX PT Hepatocyte growth factor HGF derivative for inducing angiogenesis,

PT comprises fusion protein containing HGF polypeptide and polypeptide other
PT than HGF, and polypeptide having collagen binding property connected
PT directly or through intron.

XX Example 1; SEQ ID NO 3; 51pp; Japanese.

XX This invention relates to a novel hepatocyte growth factor (HGF)
CC derivative. Specifically, it refers to a fusion protein containing an HGF
CC polypeptide and the collagen binding domain of fibronectin, where the
CC collagen binding property and the HGF activity are maintained. The
CC present invention describes this HGF derivative as a live organ
CC transplant material that can be an artificial blood vessel or a stent in
CC which the collagen material is bound on its surface in a sheet-like
CC molding. Accordingly, it is useful for inducing and enhancing
CC angiogenesis (compared to natural type HGF) in damaged tissue and thus
CC can enhance tissue regeneration. Furthermore, it exhibits vulnerary
CC activity with increased stability and collagen binding properties. This
CC polynucleotide sequence is the DNA encoding the human fibronectin
CC collagen binding domain from Ala260 to Arg599, given in an
CC exemplification of the invention.

XX SQ Sequence 1051 BP; 274 A; 262 C; 281 G; 234 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.09e-09 Length: 1051
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-581-651D-9 (1-21) x ADS19042 (1-1051)

Qy 1 ValGlyAenGlyAArgGlyGluTrpThrCysIleAlaTySerGlnLeuArgAspGlnCys 20
Db 640 GTTGGGAATGGTCGTGGGAATGGACATGCTACTTCGACGTCGAGATCAGTGC 699

Qy 21 Ile 21

Db 700 ATT 702

RESULT 2

AAA64263
ID AAA64263 standard; DNA; 1053 BP.

XX AC AAA64263;

XX DT 20-DEC-2000 (first entry)

XX cDNA encoding a modified fibronectin collagen-binding domain.

XX Fibronectin; collagen-binding domain; sustained release; gene therapy;
KW physiologically active polypeptide; topical retention;
KW tissue regeneration; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers
FT CDS 16..1047
FT /tag= a
FT /product= "fibronectin collagen-binding domain"

XX WO200049159-A1.
XX 24-AUG-2000.

XX 21-FEB-2000; 2000WO-JP000964.

XX 19-FEB-1999; 99JP-00041913.

XX 01-NOV-1999; 99JP-00311364.

XX (TERU) TERUMO CORP.

XX

PI Ishikawa T, Kitajima T;

XX WPI; 2000-565375/52.

DR P-PSDB; AAB08505.

XX Collagen-binding active polypeptide for use in an agent for enabling
PT topical retention or sustained release of a physiologically active
PT peptide or physiological activity-imparting agent comprises a fibronectin
PT peptide.

XX Example 1; Page 126-127; 135pp; English.

XX The present sequence encodes a modified fibronectin collagen-binding
CC domain. The protein is used to construct a collagen-binding
CC physiologically active polypeptide. This polypeptide comprises a peptide
CC from fibronectin ligated to a physiologically active peptide. The
CC polypeptides are used in an agent for enabling topical retention or
CC sustained release of a physiologically active peptide or physiological
CC activity-imparting agent. They may be used in gene therapy and in tissue
CC regeneration

XX SQ Sequence 1053 BP; 275 A; 261 C; 281 G; 236 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.09e-09 Length: 1053
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-581-651D-9 (1-21) x AAA64263 (1-1053)

Qy 1 ValGlyAenGlyAArgGlyGluTrpThrCysIleAlaTySerGlnLeuArgAspGlnCys 20
Db 643 GTTGGGAATGGTCGTGGGAATGGACATGCTACTTCGACGTCGAGATCAGTGC 702

Qy 21 Ile 21

Db 703 ATT 705

RESULT 3

ABL41020
ID ABL41020 standard; DNA; 1053 BP.

XX AC ABL41020;

XX DT 12-AUG-2002 (first entry)

XX Modified human fibronectin collagen-binding domain encoding DNA.

XX Osteogenesis; drug delivery system; DDS; collagen; osteopathic; human;
KW fibronectin; ds.

XX OS Homo sapiens.

XX Key Location/Qualifiers
FT CDS 16..1047
FT /tag= a
FT /product= "modified fibronectin collagen-binding domain"

XX conflict 109
XX /tag= b
XX conflict 206
XX /tag= c
XX conflict 270
XX /tag= d
XX conflict 374
XX /tag= e
XX conflict 681
XX /tag= f

XX JP2002058485-A.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 562.251 Seconds
(without alignments)
1809.796 Million cell updates/sec

Title: US-09-581-651D-9
Perfect score: 120
Sequence: 1 VGNRGWETCIAYSLRDQCI 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n_model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092223_28778/app_query.fasta_1.1834
-DB=genEmbl -QFMT=fascap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651 @CGN 1.1.9936 @runat_07112005_092223_28778 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	358	6	CQ685590 Sequence
2	120	100.0	1053	6	BD133458 Collagen
3	120	100.0	1053	6	BD133469 Hybrid po
4	120	100.0	1053	6	E63263 Collagen-bi

5	120	100.0	1224	6	E63271
6	120	100.0	1407	6	BD133465
7	120	100.0	1428	6	BD133476
8	120	100.0	1443	6	BD133466
9	120	100.0	1527	6	E63270
10	120	100.0	1560	6	BD133477
11	120	100.0	1929	6	CQ871810
12	120	100.0	1929	6	CQ871828
13	120	100.0	2147	6	BD137021
14	120	100.0	2147	6	AX003229
15	120	100.0	2147	9	HSA276395
16	120	100.0	2192	9	HSA535086
17	120	100.0	2358	6	CQ875358
18	120	100.0	2409	6	CQ731571
19	120	100.0	4321	9	HSM806214
20	120	100.0	6816	6	CQ715726
21	120	100.0	7299	9	HSM807162
22	120	100.0	7323	11	BV178397
23	120	100.0	7502	9	HSM804082
24	120	100.0	7544	9	HSM803509
25	120	100.0	7679	6	AR274901
26	120	100.0	7680	6	AR380744
27	120	100.0	7680	6	AX277596
28	120	100.0	7680	6	AX335368
29	120	100.0	7680	9	HSPFBI
30	120	100.0	7705	6	AI14133
31	120	100.0	7705	6	AR034630
32	120	100.0	7705	6	E01162
33	120	100.0	7705	6	I70110
34	120	100.0	7705	6	AR364992
35	120	100.0	7753	9	AB191261
36	120	100.0	7803	6	AR051657
37	120	100.0	7868	9	HSM806653
38	120	100.0	7951	9	HSM806901
39	120	100.0	7951	9	HSM806902
40	120	100.0	8027	6	CQ833991
41	120	100.0	8027	6	CQ875357
42	120	100.0	8030	9	HSM806170
43	120	100.0	8035	9	HSM806171
44	120	100.0	8042	9	HSM806805
45	120	100.0	8044	6	AR454662

ALIGNMENTS

RESULT 1	CQ685590	Sequence 30516 from Patent WO02070737.	358 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CQ685590	Sequence 30516 from Patent WO02070737.				
DEFINITION	CQ685590	Sequence 30516 from Patent WO02070737.				
ACCESSION	CQ685590	Sequence 30516 from Patent WO02070737.				
VERSION	CQ685590.1	GI:42213078				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1	Liew, C.C., Marshall, W.B. and Zhang, H.				
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE		Compositions and methods relating to osteoarthritis				
JOURNAL		Patent: WO 02070737-A 30516 12-SEP-2002;				
FEATURES		Chondrocyte Inc. (CA)				
source		Location/Qualifiers				
		1..358				
		/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				
ORIGIN						
Alignment Scores:						
Pred. No.:		7.59e-10				358
Score:		120.00				21
Percent Similarity:		100.00%				0
Best Local Similarity:		100.00%				0

Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-9 (1-21) x CQ685590 (1-358)

Qy 1 ValGlyAenGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
|||||
Db 199 GTTGGGAATGTCGTGGGGAATGGACATGCTACTGCTACTCGAGTCAGATCAGTGC 258

Qy 21 Ile 21
|||
Db 259 ATT 261

RESULT 2

BD133458 1053 bp DNA linear PAT 18-SEP-2002

LOCUS
DEFINITION Collagen-binding osteogenesis promotion fused protein.
ACCESSION BD133458
VERSION BD133458.1 GI:23228403
KEYWORDS JP 2002058485-A/3.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE
AUTHORS Kitajima,T. and Ishikawa,T.
TITLE Collagen-binding osteogenesis promotion fused protein
JOURNAL Patent: JP 2002058485-A 3 26-FEB-2002;
TERUMO CORP

COMMENT
OS Artificial Sequence
PN JP 2002058485-A/3
PD 26-FEB-2002
PF 16-AUG-2000 JP 2002046744
PI TAKASHI KITAJIMA,TETSUYA ISHIKAWA
PC C12N15/09,A61K38/48,A61K48/00,A61P19/08,A61P19/10,C07K14/51,
PC C07K14/78,
PC C07K19/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N5/10,A61K37/02,A61K37/00, PC

FEATURES
source
1..1053
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Alignment Scores:
Pred. No.: 2.33e-09 Length: 1053
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-9 (1-21) x BD133458 (1-1053)

Qy 1 ValGlyAenGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
|||||
Db 643 GTTGGGAATGTCGTGGGGAATGGACATGCTACTGCTACTCGAGTCAGATCAGTGC 702

Qy 21 Ile 21
|||
Db 703 ATT 705

RESULT 4

BD133458 1053 bp DNA linear PAT 27-AUG-2002

LOCUS
DEFINITION Collagen-binding physiologically active polypeptide.
ACCESSION BD133458
VERSION BD133458.1 GI:22556336
KEYWORDS JP 2001190280-A/3.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE
AUTHORS Ishikawa,T. and Kitajima,T.
TITLE Collagen-binding physiologically active polypeptide
JOURNAL Patent: JP 2001190280-A 3 17-JUL-2001;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 6.14379 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	749	2 Q7Z391	Q7Z391 homo sapien
2	110	100.0	1103	2 Q6MZFA	Q6MZFA homo sapien
3	110	100.0	2193	2 Q6MZM7	Q6MZM7 homo sapien
4	110	100.0	2240	2 Q6MDP8	Q6MDP8 homo sapien
5	110	100.0	2265	1 F1NC_BOVIN	P07589 bos taurus
6	110	100.0	2267	2 Q6BDP9	Q6BDP9 homo sapien
7	110	100.0	2296	2 Q6N0A6	Q6N0A6 homo sapien
8	110	100.0	2357	2 Q6BDT4	Q6BDT4 homo sapien
9	110	100.0	2444	2 Q6N025	Q6N025 homo sapien
10	110	100.0	2477	2 Q6MZU5	Q6MZU5 homo sapien
11	106	96.4	2386	1 F1NC_HUMAN	P02751 homo sapien
12	105	95.5	2477	1 F1NC_MOUSE	P11276 mus musculus
13	105	95.5	2477	1 F1NC_RAT	P04937 rattus norv
14	81	73.6	2481	1 F1NC_XENLA	Q91740 xenopus lae
15	81	73.6	2481	2 Q6GOA5	Q6GOA5 xenopus lae
16	75	68.2	922	2 Q3405	Q3405 brachydanio
17	75	68.2	2478	2 Q3406	Q3406 brachydanio
18	73	66.4	2478	2 Q6JAN2	Q6JAN2 brachydanio
19	57	51.8	670	2 Q6DF16	Q6DF16 xenopus tro
20	57	51.8	674	2 Q6STC6	Q6STC6 cyprinus ca
21	57	51.8	680	2 Q7T317	Q7T317 brachydanio
22	56	50.9	686	2 Q9DE15	Q9DE15 gallus gall
23	55	50.0	707	1 MM09_HUMAN	P14780 homo sapien
24	54	49.1	679	2 Q98856	Q98856 cynops pyrr
25	53	48.2	730	1 MM09_MOUSE	P41245 mus musculus
26	53	48.2	730	2 Q80XI8	Q80XI8 mus musculus
27	52	47.3	3519	1 OL56_STRAT	Q07017 streptomyce
28	52	47.3	4150	2 Q9KIV4	Q9KIV4 streptomyce
29	51	46.4	921	2 Q8BUC1	Q8BUC1 mus musculus
30	51	46.4	1307	2 Q80V33	Q80V33 mus musculus
31	51	46.4	2078	2 Q69ZN7	Q69ZN7 mus musculus

32 50 45.5 212 2 Q9T1D8
33 50 45.5 675 2 Q8QFO6
34 49 44.5 94 2 Q86MY2
35 49 44.5 458 2 Q6Y7P9
36 49 44.5 671 2 Q6PE33
37 49 44.5 671 2 Q9W7L6
38 49 44.5 690 2 Q9PVN5
39 48 43.6 94 2 Q86MY3
40 48 43.6 208 2 Q9XSD1
41 48 43.6 673 2 Q90VB3
42 48 43.6 704 1 MM09_CANFA
43 48 43.6 704 2 Q71U09
44 48 43.6 747 2 Q9S241
45 47 42.7 157 2 P94805

ALIGNMENTS

RESULT 1
Q7Z391 PRELIMINARY; PRT; 749 AA.
AC Q7Z391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Blocher H., Boecher M., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -.
DR HSSP; Q96KP7; 1PBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 85524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 100.0%; Score 110; DB 2; Length 749;
Best Local Similarity 100.0%; Pred. No. 6.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTRGNSNGALCH 20
|||||
Db 494 TDHTVLVQTRGNSNGALCH 513

RESULT 2

Q6MZFA PRELIMINARY; PRT; 1103 AA.
ID Q6MZFA

AC Q6MZFA;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219) (Fragment).

```
GN Name=DKFZp686K139; Synonym=DKFZp686F219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG The German Human cDNA Consortium;
RA Ansoerge W., Krieger S., Regler T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CAE46200.1; -.
DR GO; BX640802; CAE45885.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS0853; FN3; 4.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82F8C4CAF634AD56 CRC64;

Query Match 100.0%; Score 110; DB 2; Length 1103;
Best Local Similarity 100.0%; Pred. No. 9.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDHTVLVQTRGGNSGALCH 20
DB 491 TDHTVLVQTRGGNSGALCH 510

RESULT 3
OSMZW7
AC Q6MZW7 PRELIMINARY; PRT; 2193 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686O12165 (Fragment).
GN Name=DKFZp686O12165;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Human uterus endothel primary cell culture;
RG The German Human cDNA Consortium;
RA Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640999; CAE46002.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003962; FN_III subd.
DR InterPro; IPR003961; FN_III.

Name=DKFZp686K139; Synonym=DKFZp686F219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Human uterus endothel primary cell culture;
RG The German Human cDNA Consortium;
RA Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640999; CAE46002.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003962; FN_III subd.
DR InterPro; IPR003961; FN_III.
```

```
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 7.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 17.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 7.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 17.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 7.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS0853; FN3; 17.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 2193 AA; 240641 MW; F876B93106540EF3 CRC64;

Query Match 100.0%; Score 110; DB 2; Length 2193;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDHTVLVQTRGGNSGALCH 20
DB 149 TDHTVLVQTRGGNSGALCH 168

RESULT 4
Q68DP8
AC Q68DP8 PRELIMINARY; PRT; 2240 AA.
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686H0342.
GN Name=DKFZp686H0342;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RG The German cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749317; CAH18172.1; -.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003962; FN_III subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 10.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 1.51634 Seconds
(without alignments)
1269.066 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	2265	1 FNBO	fibronectin - bovi
2	106	96.4	2386	1 FNHU	fibronectin precu
3	105	95.5	2477	2 S14428	fibronectin precu
4	81	73.6	2481	2 A43908	fibronectin - Afri
5	55	50.0	707	1 A34458	gelatinase B (EC 3
6	53	48.2	730	1 I52580	gelatinase B (EC 3
7	53	48.2	730	2 J1456	gelatinase B (EC 3
8	52	47.3	3519	2 S43048	polyketide synthas
9	48	43.6	747	2 T36812	probable dehydroge
10	47	42.7	157	2 T44794	hypothetical prote
11	47	42.7	340	2 T32931	hypothetical prote
12	47	42.7	340	2 C87732	protein W03D8.4 [i
13	47	42.7	438	1 A57667	pop-1 protein - Ca
14	47	42.7	708	2 J43364	gelatinase B (EC 3
15	47	42.7	708	2 S62907	gelatinase B (EC 3
16	47	42.7	1079	2 A70972	probable DNA polym
17	46	41.8	493	2 T06031	hexokinase homolog
18	46	41.8	707	1 A53796	gelatinase B (EC 3
19	46	41.8	712	1 I46031	gelatinase B (EC 3
20	44	40.0	413	2 J23135	chitinase (EC 3.2.
21	44	40.0	605	2 T09892	hypothetical prote
22	44	40.0	843	2 T06068	probable proton pu
23	43	39.1	237	2 A98357	hypothetical prote
24	43	39.1	237	2 A92925	hypothetical prote
25	43	39.1	692	2 H69961	transcription regu
26	43	39.1	1430	2 T12449	hypothetical prote
27	42.5	38.6	183	2 D87614	hypothetical prote
28	42	38.2	288	2 T10477	sec13 protein - ye
29	42	38.2	801	2 H83737	glucosidase BH0704

ALIGNMENTS

RESULT 1

FNBO

fibronectin - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004

C:Accession: A26452; B21165; A23292

R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.

Eur. J. Biochem. 161, 441-453, 1986

A:Title: Complete primary structure of bovine plasma fibronectin.

A:Reference number: A26452; MUID:87054047; PMID:3780752

A:Accession: A26452

A:Molecule type: protein

A:Residues: 1-2265 <SKO>

A:Cross-references: UNIPROT:P07589

R:Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.

Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983

A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins

A:Reference number: A21165; MUID:83221567; PMID:6304699

A:Accession: B21165

A:Molecule type: mRNA

A:Residues: 2170-2265 <KOR>

A:Cross-references: GB:K00800; NID:gl63055; PIDN:AAA30521.2; PID:95713323

R:Petersen, T.E.; Thogersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sottl

Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983

A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal

A:Reference number: A23292; MUID:83117805; PMID:6218503

A:Accession: A23292

A:Molecule type: protein

A:Residues: 1-16, 'C', 18-20, 'S', 22-432:447-463;1367-1517;1567-1673;2062-2176, 'N', 2178-226

C:Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.

C:Comment: The plasma fibronectin molecule consists of two chains, which are connected b

C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib

aling, and maintenance of cell shape.

C:Comment: Plasma fibronectin is synthesized by hepatocytes.

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep

C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu

R:21-241/Domain: fibronectin type I repeat homology <1F1>

R:21-56/Domain: fibronectin type I repeat homology <1F1>

R:166-104/Domain: fibronectin type I repeat homology <1F2>

R:110-148/Domain: fibronectin type I repeat homology <1F2>

R:155-194/Domain: fibronectin type I repeat homology <1F4>

R:200-239/Domain: fibronectin type I repeat homology <1F5>

R:277-577/Domain: collagen binding <CBR>

R:277-311/Domain: fibronectin type I repeat homology <1F6>

R:329-370/Domain: fibronectin type II repeat homology <2F1>

R:389-430/Domain: fibronectin type II repeat homology <2F2>

R:439-477/Domain: fibronectin type I repeat homology <1F7>

R:487-524/Domain: fibronectin type I repeat homology <1F8>

R:530-568/Domain: fibronectin type I repeat homology <1F9>

R:578-661/Domain: fibronectin type III repeat homology <FN3A>

R:688-770/Domain: fibronectin type III repeat homology <FN3B>

R:779-860/Domain: fibronectin type III repeat homology <FN3C>

unknown protein P5
BLN protein [limp
vpr protein - huma
gas-vesicle operon
gas-vesicle operon
fibronectin - eac
tegalose-1,6-bisph
cell division prot
hypothetical prote
gelatinase A (EC 3
gelatinase A (EC 3
gelatinase A (EC 3
gelatinase A (EC 3
gelatinase A (EC 3
hypothetical prote
ribonucleoside-dip

F1875-957/Domain: fibronectin type III repeat homology <FN3D>
 F1965-1046/Domain: fibronectin type III repeat homology <FN3E>
 F1055-1134/Domain: fibronectin type III repeat homology <FN3F>
 F1142-1227/Domain: fibronectin type III repeat homology <FN3G>
 F1235-1318/Domain: fibronectin type III repeat homology <FN3H>
 F1336-1404/Domain: fibronectin type III repeat homology <FN3I>
 F1410-1517/Domain: cell attachment <CAD>
 F1416-1502/Domain: fibronectin type III repeat homology <FN3J>
 F1493-1495/Region: cell attachment (R-G-D) motif
 F1510-1592/Domain: fibronectin type III repeat homology <FN3K>
 F1600-1870/Domain: heparin binding <HB2>
 F1600-1682/Domain: fibronectin type III repeat homology <FN3L>
 F1692-1773/Domain: fibronectin type III repeat homology <FN3M>
 F1781-1863/Domain: fibronectin type III repeat homology <FN3N>
 F1970-1972/Region: cell attachment (R-G-D) motif
 F1982-2062/Domain: fibronectin type III repeat homology <FN3O>
 F1985-2216/Domain: fibrin binding <FB>
 F2085-2124/Domain: fibronectin type I repeat homology <IF10>
 F2130-2146/Domain: fibronectin type I repeat homology <IF11>
 F2174-2209/Domain: fibronectin type I repeat homology <IF12>
 F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status expirime
 F21-47-45-56-66-94-92-104-110-138-136-148-155-184-182-194-200-229-227-239-277-304-302-3
 7-2155-2167-2174-2200-2198-2209/Disulfide bonds: #status predicted
 F1999-497-511-846-976-1213-1987/Binding site: carbonylhydrate (Asn) (covalent) #status expe
 F1205-1692/Binding site: carbonylhydrate (Asn) (covalent) #status absent
 F1943-1944/Binding site: carbonylhydrate (Thr) (covalent) #status experimental
 F2246/Disulfide bonds: interchain (to 2250) #status predicted
 F2250/Disulfide bonds: interchain (to 2246) #status predicted
 F2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 110; DB 1; Length 2265;
 Best Local Similarity 100.0%; Pred. NO. 2.1e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDHTLVQTRGNSGNALCH 20
 DB 371 TDHTLVQTRGNSGNALCH 390

RESULT 2
 FNHU
 fibronectin precursor [validated] - human
 N1/Alternate names: fibronectin splice form ED-A
 C1/Species: Homo sapiens (man)
 C2/Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004
 C3/Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22
 R1/Dean, D.C.; Bowls, C.L.; Bourgeois, S.
 Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
 A1/Title: Cloning and analysis of the promoter region of the human fibronectin gene.
 A1/Reference number: A26460; MUID:87175578; PMID:3031656
 A1/Accession: A26460
 A1/Molecule type: DNA
 A1/Residues: 1-49 <DEA>
 A1/Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAA5337
 R1/Oldberg, A.; Ruoslahti, E.
 J. Biol. Chem. 261, 2113-2116, 1986
 A1/Title: Evolution of the fibronectin gene.
 A1/Reference number: A26284; MUID:86111901; PMID:3003095
 A1/Accession: A26284
 A1/Molecule type: DNA
 A1/Residues: 1447-1540 <OLD>
 A1/Cross-references: GB:M12549; NID:g182688
 A1/Note: the authors translated the codon TTC for residue 1494 as Glu
 R1/Paoletti, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.
 Nucleic Acids Res. 16, 3545-3557, 1988
 A1/Title: Sequence analysis and in vivo expression show that alternative splicing of ED-E
 A1/Reference number: S00848; MUID:882333940; PMID:3375063
 A1/Accession: S03917
 A1/Molecule type: DNA
 A1/Residues: 1594-1767 'V', 1769-1783 <PAO>
 A1/Cross-references: EMBL:X07718; NID:g14102
 A1/Note: the authors translated the codon AAC for residue 1631 as Asp

R1/Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
 FEBS Lett. 207, 287-291, 1986
 A1/Title: Donor and acceptor splice signals within an exon of the human fibronectin gene.
 A1/Reference number: A24854; MUID:87030929; PMID:3770201
 A1/Accession: A24854
 A1/Molecule type: DNA
 A1/Residues: 1992-2147 <VIB>
 A1/Cross-references: GB:X04530; NID:g31436
 R1/Gutman, A.; Yamada, K.M.; Kornblitt, A.
 FEBS Lett. 207, 145-148, 1986
 A1/Title: Human fibronectin is synthesized as a pre-propolypeptide.
 A1/Reference number: A24476; MUID:87030890; PMID:3770189
 A1/Accession: A24476
 A1/Status: not compared with conceptual translation
 A1/Molecule type: mRNA
 A1/Residues: 1-14, 'Q', 16-38 <GUT>
 R1/Kornblitt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
 EMBO J. 4, 1755-1759, 1985
 A1/Title: Primary structure of human fibronectin: differential splicing may generate at l
 A1/Reference number: A91008; MUID:85284965; PMID:2992939
 A1/Accession: A91008
 A1/Status: nucleic acid sequence not shown
 A1/Molecule type: mRNA
 A1/Residues: 32-1344, 1346-2080; 2112-2386 <KOR>
 A1/Cross-references: GB:X02761
 R1/Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Nucleic Acids Res. 12, 5853-5868, 1984
 A1/Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
 A1/Reference number: A93529; MUID:84272258; PMID:6462919
 A1/Accession: A93529
 A1/Molecule type: mRNA
 A1/Residues: 973-2080; 2112-2386 <K02>
 A1/Cross-references: GB:X00739
 R1/Oldberg, A.; Linney, E.; Ruoslahti, E.
 J. Biol. Chem. 258, 10193-10196, 1983
 A1/Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell
 A1/Reference number: A21011; MUID:83290929; PMID:6688418
 A1/Accession: A21011
 A1/Molecule type: mRNA
 A1/Residues: 1434-1537 <OL2>
 A1/Cross-references: GB:M0055; NID:g182680; PIDN:AAA52459.1; PID:g182683
 R1/Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
 Biochemistry 24, 2698-2704, 1985
 A1/Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with
 A1/Reference number: A90495; MUID:85280409; PMID:2992573
 A1/Accession: A90495
 A1/Molecule type: mRNA
 A1/Residues: 1594-2386 <BER>
 A1/Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697
 R1/Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.
 FEBS Lett. 186, 31-34, 1985
 A1/Title: Isolation and characterization of cDNA clones for human liver fibronectin.
 A1/Reference number: A22245; MUID:85231203; PMID:2989004
 A1/Accession: A22245
 A1/Molecule type: mRNA
 A1/Residues: 1948-2067 <UME>
 A1/Cross-references: GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:g182706
 A1/Accession: B22245
 A1/Molecule type: mRNA
 A1/Residues: 1975-1991; 2017-2039 <UM2>
 A1/Cross-references: GB:M27590
 R1/Sekiguchi, K.; Kios, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
 Biochemistry 25, 4936-4941, 1986
 A1/Title: Human liver fibronectin complementary DNAs: identification of two different mes
 A1/Reference number: 152394; MUID:87026578; PMID:3021206
 A1/Accession: I65273
 A1/Status: preliminary; translated from GB/EMBL/DBJ
 A1/Molecule type: mRNA
 A1/Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
 A1/Cross-references: GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:g182704
 R1/Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
 A1/Title: Isolation and characterization of cDNA clones for human and bovine fibronectin

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 6.77124 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGGNSNGALCH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	228	15	US-10-344-634-4
2	110	100.0	285	15	US-10-344-634-16
3	110	100.0	343	9	US-09-934-706-1
4	110	100.0	386	15	US-10-344-634-14
5	110	100.0	400	9	US-09-934-706-5
6	110	100.0	501	9	US-09-934-706-4
7	110	100.0	642	16	US-10-741-601-354
8	110	100.0	642	17	US-10-741-601-1066
9	110	100.0	657	16	US-10-741-601-359
10	110	100.0	657	17	US-10-741-600-1072
11	110	100.0	984	16	US-10-741-601-356

12	110	100.0	984	17	US-10-741-600-1069	Sequence 1069, Ap
13	110	100.0	1173	18	US-10-450-763-52634	Sequence 52634, A
14	110	100.0	2220	15	US-10-236-392-4	Sequence 4, Appli
15	110	100.0	2296	16	US-10-741-601-363	Sequence 363, App
16	110	100.0	2296	17	US-10-741-600-1075	Sequence 1075, Ap
17	110	100.0	2320	14	US-10-279-733-8	Sequence 8, Appli
18	110	100.0	2355	15	US-10-144-194A-104	Sequence 104, App
19	110	100.0	2355	15	US-10-447-161-3	Sequence 3, Appli
20	110	100.0	2355	16	US-10-734-564-94	Sequence 357, App
21	110	100.0	2355	16	US-10-741-601-357	Sequence 357, App
22	110	100.0	2355	16	US-10-741-601-366	Sequence 106, App
23	110	100.0	2355	16	US-10-491-566-104	Sequence 104, App
24	110	100.0	2355	17	US-10-741-600-1067	Sequence 1067, Ap
25	110	100.0	2355	17	US-10-741-600-1078	Sequence 1078, Ap
26	110	100.0	2355	17	US-10-852-335A-147	Sequence 147, App
27	110	100.0	2355	18	US-10-287-436A-436	Sequence 436, App
28	110	100.0	2355	18	US-10-287-436A-1137	Sequence 1137, Ap
29	110	100.0	2355	20	US-11-040-130-28	Sequence 28, Appli
30	110	100.0	2386	16	US-10-741-601-360	Sequence 360, App
31	110	100.0	2386	16	US-10-618-281-32	Sequence 1071, Ap
32	110	100.0	2386	17	US-10-741-600-1071	Sequence 2, Appli
33	106	96.4	2320	15	US-10-236-392-2	Sequence 64, Appli
34	106	96.4	2328	14	US-10-171-311-64	Sequence 70, Appli
35	106	96.4	2328	15	US-10-236-031B-70	Sequence 98, Appli
36	106	96.4	2328	15	US-10-374-979-98	Sequence 677, App
37	106	96.4	2328	16	US-10-182-936A-98	Sequence 677, App
38	106	96.4	2328	16	US-10-477-238A-677	Sequence 677, App
39	106	96.4	2328	16	US-10-680-287A-677	Sequence 677, App
40	106	96.4	2328	17	US-10-477-173-677	Sequence 677, App
41	106	96.4	2355	15	US-10-360-101-235	Sequence 235, App
42	106	96.4	2386	10	US-09-961-403-1	Sequence 1, Appli
43	106	96.4	2386	16	US-10-788-792-206	Sequence 206, App
44	106	96.4	2386	17	US-10-868-577A-59	Sequence 59, Appli
45	106	96.4	2386	18	US-10-485-758-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-344-634-4
; Sequence 4, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takashi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human Fibronectin
; OTHER INFORMATION: Collagen-binding Domain
US-10-344-634-4

Query Match 100.0%; Score 110; DB 15; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TDHTVLVQTRGGNSNGALCH 20
Db 113 TDHTVLVQTRGGNSNGALCH 132

```
RESULT 2
US-10-344-634-16
; Sequence 16, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takashi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Collagen-binding
; OTHER INFORMATION: Epidermal Growth Factor
US-10-344-634-16

Query Match          100.0%; Score 110; DB 15; Length 285;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTRGNSNGALCH 20
|||||
Db 113 TDHTVLVQTRGNSNGALCH 132

RESULT 3
US-09-934-706-1
; Sequence 1, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 1
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Modified Human
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN
; LOCATION: (2)..(341)
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: CONFLICT
; LOCATION: (69)
; NAME/KEY: CONFLICT
; LOCATION: (125)
US-09-934-706-1

Query Match          100.0%; Score 110; DB 9; Length 343;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTRGNSNGALCH 20
|||||
```

```
Db 113 TDHTVLVQTRGNSNGALCH 132

RESULT 4
US-10-344-634-14
; Sequence 14, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takashi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Collagen-binding
; OTHER INFORMATION: Basic Fibroblast Growth Factor
US-10-344-634-14

Query Match          100.0%; Score 110; DB 15; Length 386;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTRGNSNGALCH 20
|||||
Db 113 TDHTVLVQTRGNSNGALCH 132

RESULT 5
US-09-934-706-5
; Sequence 5, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 5
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hybrid
; OTHER INFORMATION: Polypeptide of Human Fibronectin Collagen-Binding
; OTHER INFORMATION: Domain and Human Epidermal Growth Factor
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN
; LOCATION: (2)..(341)
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: PEPTIDE
; LOCATION: (343)..(347)
; OTHER INFORMATION: /note="enterokinase recognition sequence"
; NAME/KEY: PEPTIDE
; LOCATION: (348)..(400)
; OTHER INFORMATION: /note="human epidermal growth factor"
US-09-934-706-5

Query Match          100.0%; Score 110; DB 9; Length 400;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 2.11765 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-8
Perfect score: 110
Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	96.4	2231	1 US-08-153-799-16	Sequence 16, Appl
2	106	96.4	2324	1 US-08-283-857-1	Sequence 1, Appl
3	106	96.4	2324	5 PCT-US95-09819-1	Sequence 1, Appl
4	106	96.4	2327	6 5455158-1	Patent No. 5455158
5	106	96.4	2327	6 5455158-1	Patent No. 5455158
6	106	96.4	2386	2 US-09-016-366A-12	Sequence 12, Appl
7	106	96.4	2386	4 US-09-961-403-1	Sequence 1, Appl
8	106	96.4	2446	2 US-08-551-356-2	Sequence 2, Appl
9	106	96.4	2446	5 PCT-US93-12687-2	Sequence 2, Appl
10	87	79.1	65	2 US-08-982-597A-19	Sequence 19, Appl
11	87	79.1	65	3 US-09-136-218-19	Sequence 19, Appl
12	55	50.0	707	3 US-08-704-711A-19	Sequence 19, Appl
13	55	50.0	707	3 US-09-521-220-19	Sequence 19, Appl
14	55	50.0	707	3 US-09-391-104-20	Sequence 20, Appl
15	55	50.0	707	4 US-09-949-016-6575	Sequence 20, Appl
16	55	50.0	708	3 US-08-448-489-16	Sequence 16, Appl
17	55	50.0	708	4 US-09-689-730-16	Sequence 16, Appl
18	55	50.0	713	4 US-09-949-016-10629	Sequence 16, Appl
19	53	48.2	60	2 US-08-982-597A-18	Sequence 18, Appl
20	53	48.2	60	3 US-09-136-218-18	Sequence 18, Appl
21	52	47.3	3519	3 US-09-428-517-4	Sequence 4, Appl
22	52	47.3	4150	3 US-09-428-517-2	Sequence 2, Appl
23	48	43.6	13	2 US-08-982-597A-8	Sequence 8, Appl
24	48	43.6	13	3 US-09-136-218-8	Sequence 8, Appl
25	48	43.6	14	2 US-08-982-597A-9	Sequence 9, Appl
26	48	43.6	14	3 US-09-136-218-9	Sequence 9, Appl
27	48	43.6	15	2 US-08-982-597A-1	Sequence 1, Appl

Sequence 1, Appl
Sequence 33829, A
Sequence 49046, A
Sequence 14, Appl
Sequence 14, Appl
Sequence 44610, A
Sequence 1, Appl
Sequence 6520, Ap
Sequence 8335, Ap
Sequence 7, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 59273, A
Sequence 374, App
Sequence 4801, Ap

ALIGNMENTS

RESULT 1
US-08-153-799-16
; Sequence 16, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..2231
; OTHER INFORMATION: /note= "Human fibronectin"
US-08-153-799-16

Query Match          96.4%; Score 106; DB 1; Length 2231;
Best Local Similarity 95.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTGGNSGALCH 20
Db 371 TDHTVLVQTGGNSGALCH 390

RESULT 2
US-08-283-857-1
; Sequence 1, Application US/08283857
; Patent No. 5792742
; GENERAL INFORMATION:
; APPLICANT: GOLD, Leslie I.
; APPLICANT: ROSTAGNO, Agueda A.
; APPLICANT: BARON, Martin
; APPLICANT: CAMPBELL, Iain D.
; APPLICANT: WILLIAMS, Michael, J.
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USERS THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSER: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/283,857
; FILING DATE: 01-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/714,134
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-283-857-1

Query Match          96.4%; Score 106; DB 1; Length 2324;
Best Local Similarity 95.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTGGNSGALCH 20
Db 371 TDHTVLVQTGGNSGALCH 390

RESULT 3
PCT-US95-09819-1
; Sequence 1, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-09819-1

Query Match          96.4%; Score 106; DB 5; Length 2324;
Best Local Similarity 95.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTGGNSGALCH 20
Db 371 TDHTVLVQTGGNSGALCH 390

RESULT 4
5455158-1
; Patent No. 5455158
; APPLICANT: VOGEL, TIKVA; LEVANON, AVIGDOR; WERBER, MOSHE M.;
; GUY, RACHEL; PANET, AMOS
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,241
; FILING DATE: 04-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 526,397
; FILING DATE: 21-MAY-1990
; APPLICATION NUMBER: 345,952
; FILING DATE: 28-APR-1989
; APPLICATION NUMBER: 291,951
; FILING DATE: 29-DEC-1988
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 7.39869 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	20	2	AAY28913
2	110	100.0	21	2	AAY28905
3	110	100.0	228	5	ABB06747 Human fib
4	110	100.0	285	5	ABB06751 Collagen-
5	110	100.0	343	3	ABB08505 Amino aci
6	110	100.0	343	5	ABB07961 Modified
7	110	100.0	386	5	ABB06750 Collagen-
8	110	100.0	400	3	ABB08509 Hybird of
9	110	100.0	461	5	ABB07964 Human fib
10	110	100.0	473	5	ABB07965 Human fib
11	110	100.0	501	3	ABB08508 Hybird of
12	110	100.0	642	8	ADQ39403 Human myo
13	110	100.0	642	8	ADR67316 Human bla
14	110	100.0	642	8	ADSL17489 Amino aci
15	110	100.0	642	8	ADR97658 Human fib
16	110	100.0	657	8	ADQ39409 Human myo
17	110	100.0	660	2	AAY28901 Human mig
18	110	100.0	951	8	ADSL19044 Chimeric
19	110	100.0	984	8	ADQ39406 Human myo
20	110	100.0	1173	4	ABG22275 Novel hum
21	110	100.0	1179	8	ADP75952 Human min
22	110	100.0	1223	8	ADP30365 Human sec
23	110	100.0	1359	8	ADP75957 Human leu
24	110	100.0	2182	8	ADR66462 Human pro
25	110	100.0	2182	8	ADR66120 Human pro

26	110	100.0	2220	6	ABO01289	AbO01289 Human pro
27	110	100.0	2265	4	AAAM38647	AAm38647 Human pol
28	110	100.0	2265	8	ADSL17498	ADSL17498 Amino aci
29	110	100.0	2265	8	ADR97667	ADR97667 Bovine fi
30	110	100.0	2265	6	ABR40124	ABr40124 Human cel
31	110	100.0	2296	8	ADQ39412	ADq39412 Human myo
32	110	100.0	2320	6	AAE37107	AAE37107 Human fib
33	110	100.0	2320	6	ABR42588	ABr42588 Human fib
34	110	100.0	2330	4	AAAM38646	AAm38646 Human pol
35	110	100.0	2355	4	AAAM38649	AAm38649 Human pol
36	110	100.0	2355	6	ABR58335	ABr58335 NM_00202
37	110	100.0	2355	7	ADP65196	ADp65196 Human fib
38	110	100.0	2355	8	ADG89560	ADg89560 Human fib
39	110	100.0	2355	8	ADOS5175	ADOS5175 Protein #
40	110	100.0	2355	8	ADQ26085	ADq26085 Fibronect
41	110	100.0	2355	8	ADQ29668	ADq29668 Human col
42	110	100.0	2355	8	ADQ39415	ADq39415 Human myo
43	110	100.0	2355	8	ADQ39404	ADq39404 Human myo
44	110	100.0	2355	8	ADR67315	ADR67315 Human bla
45	110	100.0	2386	4	AAAM38648	AAm38648 Human pol

ALIGNMENTS

RESULT 1

AAAY28913
ID AAY28913 standard; peptide; 20 AA.

AC AAY28913;

DT 21-SEP-1999 (first entry)

DE MSF 1-alpha peptide epitope.

KW Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; epitope; fibronectin.

OS Synthetic.

OS Homo sapiens.

PN WO9931233-A1.

XX 24-JUN-1999.

PF 15-DEC-1998; 98WO-GB003766.

PR 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI; 1999-430039/36.

XX Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX Claim 29; Page 61; 86pp; English.

XX The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. Sequences AAY28912-913 represent peptide epitopes of MSF against which monoclonal antibodies can be raised

SQ Sequence 20 AA;

Query Match 100.0%; Score 110; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 5.3e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TDHTVLVQTRGNSNGALCH 20
Db 1 TDHTVLVQTRGNSNGALCH 20

RESULT 2
AAV28905
ID AAV28905 standard; peptide; 21 AA.
XX
AC AAV28905;
XX
DT 21-SEP-1999 (first entry)
XX
DE MSF 1-alpha peptide epitope.
XX
KW Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; epitope; fibronectin.
XX
OS Synthetic.
OS Homo sapiens.
PN WO9931233-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-GB003766.
XX
PR 16-DEC-1997; 97GB-00026539.
XX
PA (UYDU-) UNIV DUNDEE.
XX
PI Schor SL, Schor AM;
XX
DR WPI; 1999-430039/36.
XX
PT Proteins with cell migration stimulatory activity used in treating wound
PT and preventing scarring.
XX
PS Claim 18; Page 57; 86pp; English.
XX
CC The invention provides a human migration stimulatory factor (MSF)
CC protein. Host cells containing a replicable vector comprising the MSF
CC encoding nucleic acid can be used for the recombinant production of the
CC protein. The polypeptide can be used for modulating cell migration,
CC healing a wound and for preventing scarring. Sequences AAV28902-906
CC represent peptide epitopes of MSF against which monoclonal antibodies
CC that are specific to MSF and do not cross-react with fibronectin are
CC raised
XX
SQ Sequence 21 AA;
Query Match 100.0%; Score 110; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.6e-11; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 TDHTVLVQTRGNSNGALCH 20
Db 2 TDHTVLVQTRGNSNGALCH 21

RESULT 3
ABB06747
ID ABB06747 standard; protein; 228 AA.
XX
AC ABB06747;
XX
DT 13-JUN-2002 (first entry)
XX
DE Human fibronectin collagen-binding domain protein SEQ ID NO:4.
XX
KW Human; fibronectin; collagen-binding domain; collagen-binding hybrid;
KW basic fibroblast growth factor; enterokinase recognition site; collagen;
KW basic fibroblast growth factor; drug delivery system; tissue regeneration.

epidermal growth factor; drug delivery system; tissue regeneration.

KW 1 TDHTVLVQTRGNSNGALCH 20
XX
OS Homo sapiens.
PN WO200214505-A1.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-JP007036.
XX
PR 15-AUG-2000; 2000JP-00246341.
XX
PA (TERU) TERUMO CORP.
XX
PI Ishikawa T, Kitajima T;
XX
DR WPI; 2002-257605/30.
DR N-PSDB; ABL50261.
XX
PT Collagen-binding hybrid polypeptide, useful in drug delivery system for
PT functional polypeptides, formulating into complex with collagen to give
PT functionally modified collagen matrix as biomaterial for tissue
PT regeneration.
XX
PS Example 1; Page 61-63; 86pp; Japanese.
XX
CC The present invention describes a hybrid polypeptide comprising a
CC collagen-binding domain composed of an amino acid sequence ranging from
CC Ala at position 260 to Arg at position 484 of human fibronectin or a
CC similar amino acid sequence but with some amino acids deleted,
CC substituted, inserted or added, and a functional polypeptide linked
CC together. The present invention also describes: (1) a biomaterial
CC containing a functional polypeptide-modified collagen obtained by
CC formulating a polypeptide originated from the hybrid polypeptide with
CC collagen into a complex; (2) a gene encoding the hybrid polypeptide; and
CC (3) a transformant containing the gene. The hybrid polypeptide is useful
CC in drug delivery system for functional polypeptides, formulating into
CC complex with collagen to give functionally modified collagen matrix as
CC biomaterial for tissue regeneration. The hybrid polypeptide has superior
CC activity and stability in the body for a long period of time, with
CC localisability and sustained-releasability. The present sequence
CC represents the human fibronectin collagen-binding domain, which is used
CC in an example from the present invention
XX
SQ Sequence 228 AA;
Query Match 100.0%; Score 110; DB 5; Length 228;
Best Local Similarity 100.0%; Pred. No. 9.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDHTVLVQTRGNSNGALCH 20
Db 113 TDHTVLVQTRGNSNGALCH 132

RESULT 4
ABB06751
ID ABB06751 standard; protein; 285 AA.
XX
AC ABB06751;
XX
DT 13-JUN-2002 (first entry)
XX
DE Collagen-binding epidermal growth factor protein SEQ ID NO:16.
XX
KW Human; fibronectin; collagen-binding domain; collagen-binding hybrid;
KW basic fibroblast growth factor; enterokinase recognition site; collagen;
KW epidermal growth factor; drug delivery system; tissue regeneration.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200214505-A1.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 415.948 Seconds
(without alignments)
1830.243 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgm2_1/USPTO-spool_h/US0581651/runat_07112005_092223_28789/app_query.fasta_1.1834
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0581651 @CGN 1 1 8180 @runat_07112005_092223_28789 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c	1	110	100.0	263	2 B699322 RC3-NN006
	2	110	100.0	310	1 AA375384 EST87654
	3	110	100.0	373	1 AA376281 EST88736
	4	110	100.0	385	1 AL589418 DKF2p451P
c	5	110	100.0	392	2 BE079849 RC6-BT062
	6	110	100.0	400	2 BE122858 14_08 Hum
c	7	110	100.0	411	5 BX642740 DKF2p781G
	8	110	100.0	458	1 AJ661698 AJ661698
	9	110	100.0	471	4 BG007031 RC1-GN023

10	110	100.0	471	5	BX495442
11	110	100.0	478	5	BX495500
12	110	100.0	480	7	KN419466
13	110	100.0	484	4	BF986449
c	14	110	528	7	CF793207
	15	110	538	1	AJ685866
c	16	110	542	5	BQ340573
	17	110	558	7	KN419467
18	110	100.0	560	4	BG897886
19	110	100.0	605	7	KN481619
20	110	100.0	618	7	KN482377
21	110	100.0	620	7	KN481412
22	110	100.0	622	7	KN419625
23	110	100.0	632	7	KN482351
24	110	100.0	645	7	KN419499
25	110	100.0	653	7	KN419474
26	110	100.0	655	7	KN419546
27	110	100.0	657	7	KN419503
28	110	100.0	664	6	CB471201
29	110	100.0	673	1	AU140450
30	110	100.0	695	7	KN419477
31	110	100.0	702	7	KN419472
32	110	100.0	705	1	AU140899
33	110	100.0	723	7	KN419622
c	34	110	725	5	BQ574857
	35	110	730	1	AU140834
36	110	100.0	731	1	AU140526
37	110	100.0	748	1	AU140948
38	110	100.0	748	1	AU140991
39	110	100.0	758	1	AU140556
40	110	100.0	766	7	KN419594
c	41	110	780	7	KN162843
	42	110	781	7	KN161442
43	110	100.0	796	1	AU140814
44	110	100.0	816	1	AU141008
c	45	110	822	1	AA788933

ALIGNMENTS

RESULT 1
BE699322/c

LOCUS RC3-NN0064-150600-022-all NN0064

DEFINITION RC3-NN0064-150600-022-all NN0064 Homo sapiens CDNA, mRNA sequence.

ACCESSION BE699322

VERSION BE699322.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 263)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brundsen,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

263 bp mRNA linear EST 12-SEP-2000

RC3-NN0064-150600-022-all NN0064 Homo sapiens CDNA, mRNA sequence.

BE699322.1 GI:10087064

EST.

Homo sapiens (human)

Homo sapiens

1 (bases 1 to 263)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brundsen,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=6t2-RC3-NN0064-150
600-022-all&3=2000-06-15&4=1)

Seq primer: puc 18 forward
High quality sequence stop: 263.

FEATURES

source

Location/Qualifiers
1..263
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN0064"
/note="Organ: nervous normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Alignment Scores:
Pred. No.: 2,466-08 Length: 263
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-581-651D-8 (1-20) x BE699322 (1-263)

QY 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
|||||
Db 129 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAAATTCGAATGGTGCCTTGTGCCAC 70
|||||

RESULT 2

AA375384

LOCUS EST87654 HSC172 cells II Homo sapiens cDNA 5' end similar to
DEFINITION fibronectin, mRNA sequence.

ACCESSION AA375384

VERSION AA375384.1 GI:2027703

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE
AUTHORS
1 (bases 1 to 310)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

7566098

Other ESTs: THC167580

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 301869056

Fax: 3018699423

Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tadb/hgi/hgi.html)

Seq primer: M13 Reverse

Location/Qualifiers

1..310

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="ATCC (inhost):179810"

/db_xref="taxon:9606"

/cell_type="fibroblast"

/cell_line="HSC172 (60PDL)"

/dev_stage="fetal"

/clone_lib="HSC172 cells II"

/note="Organ: lung; Vector: pBlueScript SK-; Site_1:

ECOR1; Site_2: XhoI"

ORIGIN

Alignment Scores:
Pred. No.: 2,986-08 Length: 310
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-581-651D-8 (1-20) x AA375384 (1-310)

QY 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
|||||
Db 19 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAAATTCGAATGGTGCCTTGTGCCAC 78
|||||

RESULT 3

AA376281

LOCUS

DEFINITION EST88736 HSC172 cells II Homo sapiens cDNA 5' end similar to
fibronectin, mRNA sequence.

ACCESSION AA376281

VERSION AA376281.1 GI:2028652

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE
AUTHORS
1 (bases 1 to 373)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

7566098

Other ESTs: THC167580

Contact: Kerlavage, AR

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 102.693 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-8
Perfect score: 110
Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1834
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALICNA=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPSZE=500 -MINLEN=0
-NCPU=2000000000 -USER=US09581651@cgn1_1.1041 @runat_07112005_092226_28867
-NCPU=6 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq2.*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	310	9	US-09-765-231A-1
2	110	100.0	708	19	US-10-344-634-3
3	110	100.0	731	15	US-10-198-846-8918
4	110	100.0	879	19	US-10-344-634-15
5	110	100.0	1053	9	US-09-934-706-8
6	110	100.0	1182	19	US-10-344-634-13
7	110	100.0	1189	15	US-10-198-846-11556
8	110	100.0	1224	9	US-09-934-706-16
9	110	100.0	1527	9	US-09-934-706-15
10	110	100.0	2127	17	US-10-210-120-49
11	110	100.0	2127	22	US-10-956-157-4288
12	110	100.0	2127	24	US-10-909-035-49
13	110	100.0	2443	20	US-10-741-601-70
14	110	100.0	2443	22	US-10-741-600-238
15	110	100.0	2488	20	US-10-741-601-75
16	110	100.0	2488	22	US-10-741-600-244
17	110	100.0	3522	24	US-10-450-763-22266
18	110	100.0	6510	20	US-10-741-601-72
19	110	100.0	6510	22	US-10-741-600-241
20	110	100.0	7361	19	US-10-236-392-3
21	110	100.0	7677	22	US-10-956-157-4995
22	110	100.0	7795	16	US-10-084-817-2
23	110	100.0	7823	20	US-10-741-601-77
24	110	100.0	7823	22	US-10-741-600-245
25	110	100.0	7848	20	US-10-741-601-78
26	110	100.0	7848	22	US-10-741-600-246
27	110	100.0	7867	14	US-10-098-841-6
28	110	100.0	7935	22	US-10-741-601-74
29	110	100.0	7935	22	US-10-741-600-240
30	110	100.0	7959	20	US-10-741-601-81
31	110	100.0	7959	22	US-10-741-600-249
32	110	100.0	8013	20	US-10-741-601-71
33	110	100.0	8013	22	US-10-741-600-242
34	110	100.0	8027	18	US-10-447-161-8
35	110	100.0	8027	20	US-10-734-564-27
36	110	100.0	8027	22	US-10-852-335A-53
37	110	100.0	8027	24	US-10-287-436A-81
38	110	100.0	8044	17	US-10-240-965-121
39	110	100.0	8044	24	US-10-765-700-135
40	110	100.0	8062	14	US-10-098-841-5
41	110	100.0	8137	14	US-10-098-841-8
42	110	100.0	8155	20	US-10-741-601-79
43	110	100.0	8155	22	US-10-741-600-247
44	110	100.0	8226	20	US-10-741-601-69
45	110	100.0	8226	22	US-10-741-600-237

ALIGNMENTS

RESULT 1

US-09-765-231A-1
; Sequence 1, Application US/09765231A
; Patent No. US20020119452A1
; GENERAL INFORMATION:
; APPLICANT: Searle/Monsanto
; APPLICANT: Phlippard, Deborah
; APPLICANT: Vaseanthakum, Geetha
; APPLICANT: Dotson, Stanton
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,
; TITLE OF INVENTION: vectors, and cells
; FILE REFERENCE: SO-3221 PR
; CURRENT APPLICATION NUMBER: US/09/765.231A

```
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 82
; SEQ ID NO 1
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-231A-1

Alignment Scores:
Pred. No.: 1.58e-10 Length: 310
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-581-651D-8 (1-20) x US-09-765-231A-1 (1-310)
Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 19 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAATTCCAATGGTGCCTTGTGCCAC 78

RESULT 2
US-10-344-634-3
; Sequence 3, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human Fibronectin
; OTHER INFORMATION: Collagen-binding Domain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)..(693)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (109)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (206)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (270)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (374)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (681)
US-10-344-634-3

Alignment Scores:
Pred. No.: 4.11e-10 Length: 708
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-09-581-651D-8 (1-20) x US-10-344-634-3 (1-708)
Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 352 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAATTCCAATGGTGCCTTGTGCCAC 411

RESULT 3
US-10-198-846-8918
; Sequence 8918, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8918
; LENGTH: 731
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 3, 5, 6, 7, 12, 716
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-8918

Alignment Scores:
Pred. No.: 4.26e-10 Length: 731
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-581-651D-8 (1-20) x US-10-198-846-8918 (1-731)
Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 481 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAATTCCAATGGTGCCTTGTGCCAC 540

RESULT 4
US-10-344-634-15
; Sequence 15, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takashi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 879
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Collagen-binding
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:49:16 ; Search time 20.1307 Seconds
(without alignments)
1625.652 Million cell updates/sec

Title: US-09-581-651D-8
Perfect score: 110
Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool.h/US09581651/runat_07112005_092224_28801/app_query.fasta_1.1834
-DB=Issued_Patents_NA -QMT=fastcap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651 @CGN 1 1 187 @runat_07112005_092224_28801 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	8044	4	US-09-566-921-135
2	106	96.4	7679	4	Sequence 135, App
3	106	96.4	7680	4	Sequence 38, Appl
4	106	96.4	7680	5	Sequence 1289, Ap
5	106	96.4	7705	1	PCT-US95-09819-6
6	106	96.4	7705	1	Sequence 6, Appl
7	106	96.4	7705	2	Sequence 16, Appl
8	106	96.4	7705	6	Sequence 16, Appl
9	106	96.4	7803	2	Patent No. 5455158
10	106	96.4	7803	5	Patent No. 5455158
11	105	95.5	986	1	Sequence 1, Appl
12	105	95.5	986	1	Sequence 8, Appl

13	55	50.0	2109	4	US-09-799-451-345	Sequence 345, App
14	55	50.0	2334	1	US-08-457-304A-33	Sequence 33, Appl
15	55	50.0	2334	1	US-08-456-701A-33	Sequence 33, Appl
16	55	50.0	2334	3	US-08-684-932A-33	Sequence 33, Appl
17	55	50.0	2334	4	US-09-023-655-996	Sequence 996, App
18	55	50.0	2334	4	US-09-949-016-704	Sequence 704, App
19	55	50.0	2335	4	US-09-799-451-346	Sequence 346, App
20	55	50.0	2335	4	US-09-949-016-4758	Sequence 4758, Ap
c 21	53	48.2	4810	3	US-09-596-824-5	Sequence 5, Appl
c 22	53	48.2	4810	3	US-09-885-329-5	Sequence 1, Appl
23	52	47.3	50937	3	US-09-428-517-1	Sequence 1, Appl
24	51	46.4	11665	4	US-09-949-016-12446	Sequence 12446, A
25	51	46.4	11665	4	US-09-949-016-16500	Sequence 16500, A
c 26	48	43.6	10445	4	US-09-949-016-12311	Sequence 12311, A
c 27	48	43.6	10445	4	US-09-949-016-12927	Sequence 12927, A
c 28	48	43.6	58844	4	US-09-949-016-13769	Sequence 13769, A
c 29	47	42.7	1164	4	US-09-902-540-350	Sequence 350, App
c 30	47	42.7	1434	4	US-09-248-796A-1775	Sequence 1775, Ap
c 31	47	42.7	12157	4	US-09-657-960-2	Sequence 2, Appl
c 32	47	42.7	10302	4	US-09-657-960-1	Sequence 1, Appl
c 33	47	42.7	25356	4	US-09-976-594-750	Sequence 750, App
c 34	47	42.7	106746	4	US-09-326-402C-1	Sequence 1, Appl
c 35	47	42.7	106746	4	US-09-326-402C-12	Sequence 12, Appl
c 36	47	42.7	162465	4	US-09-949-016-14264	Sequence 14264, A
c 37	47	42.7	4403765	3	US-09-103-840A-2	Sequence 2, Appl
c 38	46	41.8	204	4	US-09-103-840A-1	Sequence 1, Appl
c 39	46	41.8	204	4	US-09-252-991A-9996	Sequence 9996, Ap
c 40	46	41.8	422	4	US-09-270-767-6921	Sequence 6921, Ap
c 41	46	41.8	422	4	US-09-270-767-22203	Sequence 22203, A
c 42	46	41.8	601	4	US-09-949-016-32909	Sequence 32909, A
c 43	46	41.8	601	4	US-09-949-016-67287	Sequence 67287, A
c 44	46	41.8	601	4	US-09-949-016-67288	Sequence 67288, A
c 45	46	41.8	1560	4	US-09-724-678D-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-566-921-135
; Sequence 135, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 135
; LENGTH: 8044
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 427813.14
US-09-566-921-135

Alignment Scores:
Pred. No.: 8.65e-09 Length: 8044
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-581-651D-8 (1-20) x US-09-566-921-135 (1-8044)

Qy 1 ThraspHisThrValLeuValGlnThrArgGlycysenSerArgGlyAlaLeuCyehis 20
|||
Db 1472 ACAGACCACACTGTTTGGTTTCAGACTCGAGGAGGAATTCCAATGGTCTTGTCAC 1531

RESULT 2
US-09-220-132-38
; Sequence 38, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-38

Alignment Scores:
Pred. No.: 4.14e-08 Length: 7679
Score: 106.00 Matches: 19
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.00% Mismatches: 0
Query Match: 96.36% Indels: 0
DB: Gaps: 0

US-09-581-651D-8 (1-20) x US-09-220-132-38 (1-7679)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAsnGlyAlaLeuCyHis 20
Db 1124 ACAGACCACACTGTTTGGTTCAGACTCAAGGAGGAAATTCCAATGGTGGCTTGTGCCAC 1183

RESULT 3
US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; EXPRESSION
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g31396
US-09-023-655-1289

Alignment Scores:
Pred. No.: 4.14e-08 Length: 7680
Score: 106.00 Matches: 19
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.00% Mismatches: 0
Query Match: 96.36% Indels: 0
DB: Gaps: 0

US-09-581-651D-8 (1-20) x US-09-023-655-1289 (1-7680)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAsnGlyAlaLeuCyHis 20
Db 1124 ACAGACCACACTGTTTGGTTCAGACTCAAGGAGGAAATTCCAATGGTGGCTTGTGCCAC 1183

RESULT 4
PCT-US95-09819-6
; Sequence 6, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US95-09819-6
Alignment Scores:
Pred. No.: 4.14e-08 Length: 7680

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:30:56 ; Search time 65.5686 Seconds
(without alignments)
1805.663 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool_h/US09581651/runat_07112005_092222_28772/app_query.fasta_1.1834
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THRM=ext -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651 -CGCN 1 1 1052 @runat_07112005_092222_28772 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Datatabase : N Geneseq 16Dec04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	110	100.0	310	AAH23071	Aah23071 Osteoarth
2	110	100.0	706	ADSL19041	Adsl19041 Human fib
3	110	100.0	708	ABL50261	Abi50261 Human fib
4	110	100.0	731	ACN87768	Acn87768 Breast ca
5	110	100.0	741	AAL20536	Aal20536 Human bre

6	110	100.0	879	6	ABL50269	Abi50269 Collagen-
7	110	100.0	1051	13	ADSL19042	Adsl19042 Human fib
8	110	100.0	1053	3	AAA64263	Aaa64263 cDNA enco
9	110	100.0	1053	6	ABL41020	Abi41020 Modified
10	110	100.0	1182	6	ABL50268	Abi50268 Collagen-
c 11	110	100.0	1189	11	ACN90406	Acn90406 Breast ca
12	110	100.0	1224	3	AAA64271	Aaa64271 Fibronect
13	110	100.0	1407	6	ABL41027	Abi41027 Human fib
14	110	100.0	1443	6	ABL41028	Abi41028 Human fib
15	110	100.0	1527	3	AAA64270	Aaa64270 Fibronect
16	110	100.0	1929	13	ADSL17488	Adsl17488 Nucleotid
17	110	100.0	1929	13	ADR97657	Adr97657 Human fib
18	110	100.0	2127	10	ADD18477	Add18477 Human pro
19	110	100.0	2147	2	AAX81299	Aax81299 Human mlg
20	110	100.0	2358	13	ADR67201	Adr67201 Human bla
21	110	100.0	2443	13	ADQ38575	Adq38575 Human SNP
22	110	100.0	2488	13	ADQ38581	Adq38581 Human SNP
23	110	100.0	2856	13	ADSL19043	Adsl19043 Chimeric
24	110	100.0	3522	5	AAS86462	Aas86462 DNA encod
25	110	100.0	3540	12	ADP75954	Adp75954 Human min
26	110	100.0	3669	12	ADP29134	Adp29134 Human sec
27	110	100.0	4080	12	ADP75958	Adp75958 Human leu
28	110	100.0	6510	13	ADQ38578	Adq38578 Human SNP
29	110	100.0	6816	10	ABX74443	Abx74443 Human CDN
30	110	100.0	7242	13	ADR66637	Adr66637 Human pro
31	110	100.0	7242	13	ADR65953	Adr65953 Human pro
32	110	100.0	7361	9	ACD06170	Acc06170 Human cDN
33	110	100.0	7550	8	ACC00412	Acc00412 Human cel
34	110	100.0	7677	13	ACN37820	Acn37820 Tumour-ag
35	110	100.0	7795	10	ADJ56196	Adj56196 Zebrafish
36	110	100.0	7823	13	ADQ38582	Adq38582 Human SNP
37	110	100.0	7848	13	ADQ38583	Adq38583 Human SNP
38	110	100.0	7867	4	AAI57803	Aai57803 Human pol
39	110	100.0	7935	13	ADQ38577	Adq38577 Human SNP
40	110	100.0	7959	13	ADQ38586	Adq38586 Human SNP
41	110	100.0	8013	13	ADQ38579	Adq38579 Human SNP
42	110	100.0	8027	11	ADP64998	Adp64998 Human fib
43	110	100.0	8027	12	ADG89565	Adg89565 Human fib
44	110	100.0	8027	12	ADQ29601	Adq29601 Human col
45	110	100.0	8027	13	ADR67200	Adr67200 Human bla

ALIGNMENTS

RESULT 1
AAH23071
ID AAH23071 standard; DNA; 310 BP.
XX
AC AAH23071;
DT 17-SEP-2001 (first entry)
DE Osteoarthritis tissue-derived nucleic acid sequence #1.
XX
KW Osteoarthritis; infectious disorder; Crohn's disease; sepsis; human;
XX wound healing; osteopathic; anti-arthritis; anti-inflammatory; vulnary;
XX antibacterial; antiallergic; ds.
OS Homo sapiens.
XX
PN WO200153531-A2.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US0000016.
XX
PR 18-JAN-2000; 2000US-0176523P.
XX
(PHAA) PHARMACIA CORP.
XX
PI Phippard D, Vasanthakamur G, Dotson S, Ma X;
XX
DR WPI; 2001-451914/48.

XX Substantially purified protein, polypeptide or their fragments, used to
PT identify a biologically active compound or composition and treat
PT mammalian osteoarthritis.
XX Claim 1; Page 90; 144pp; English.
XX Sequences AAH23071-23152 represent nucleic acid sequences derived from
CC osteoarthritis tissues. The sequences are useful as probes and for the
CC diagnosis or prognosis of mammalian osteoarthritis. The polynucleotides
CC and polypeptides of the invention are useful for generating diagnostic
CC reagents, as targets for small molecule drug development, generation of
CC therapeutics, and cloning genes. Specific antibodies are used to generate
CC enzyme linked immunosorbent assays for detection of osteoarthritis. The
CC invented molecules can be used to treat osteoarthritis or to analyse the
CC disease-modifying activity of osteoarthritis drugs. Other disorders
CC treatable using the nucleic acid sequences include atopic, inflammatory
CC and infectious disorders e.g. Crohn's disease and sepsis, and wound
CC healing
XX Sequence 310 BP; 87 A; 75 C; 77 G; 71 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.68e-09 Length: 310
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-581-651D-8 (1-20) x AAH23071 (1-310)
QY 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCyHis 20
DB 19 ACAGACCACACTGTTTGGTTTCAGACTCGAGGAGGAAATTCCAATGGTGCTTGTGCCAC 78
RESULT 2
ADSI19041
ID ADSI19041 standard; DNA; 706 BP.
XX
AC ADSI19041;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human fibronectin collagen binding domain DNA SeqID 2.
XX
KW human; ds; hepatocyte growth factor; HGF; collagen binding domain;
KW fibronectin; artificial blood vessel; stent; tissue regeneration;
KW vulnerable.
XX
OS Homo sapiens.
XX
PN JP2004269423-A.
XX
PD 30-SEP-2004.
XX
PF 07-MAR-2003; 2003JP-00062169.
XX
PR 07-MAR-2003; 2003JP-00062169.
XX
PA (TERU) TERUMO CORP.
XX
XX WPI; 2004-682709/67.
XX
PT Hepatocyte growth factor HGF derivative for inducing angiogenesis,
PT comprises fusion protein containing HGF polypeptide and polypeptide other
PT than HGF, and polypeptide having collagen binding property connected
PT directly or through intron.
XX
XX Example 1; SEQ ID NO 2; 51pp; Japanese.
XX
CC This invention relates to a novel hepatocyte growth factor (HGF)
CC derivative. Specifically, it refers to a fusion protein containing an HGF

CC polypeptide and the collagen binding domain of fibronectin, where the
CC collagen binding property and the HGF activity are maintained. The
CC present invention describes this HGF derivative as a live organ
CC transplant material that can be an artificial blood vessel or a stent in
CC which the collagen material is bound on its surface in a sheet-like
CC molding. Accordingly, it is useful for inducing and enhancing
CC angiogenesis (compared to natural type HGF) in damaged tissue and thus
CC can enhance tissue regeneration. Furthermore, it exhibits vulnerary
CC activity with increased stability and collagen binding properties. This
CC polynucleotide sequence is the DNA encoding the human fibronectin
CC collagen binding domain from Ala260 to Ala484, given in an
CC exemplification of the invention.
XX
SQ Sequence 706 BP; 187 A; 181 C; 187 G; 151 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4.6e-09 Length: 706
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-581-651D-8 (1-20) x ADSI19041 (1-706)
QY 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCyHis 20
DB 349 ACAGACCACACTGTTTGGTTTCAGACTCGAGGAGGAAATTCCAATGGTGCTTGTGCCAC 408
RESULT 3
ABL50261
ID ABL50261 standard; cDNA; 708 BP.
XX
AC ABL50261;
XX
DT 13-JUN-2002 (first entry)
XX
DE Human fibronectin collagen-binding domain encoding cDNA SEQ ID NO:3.
XX
KW Human; fibronectin; collagen-binding domain; collagen-binding hybrid;
KW basic fibroblast growth factor; enterokinase recognition site; collagen;
KW epidermal growth factor; drug delivery system; tissue regeneration; Gene;
KW ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 16..702
FT /*tag= a
FT /product= "human fibronectin collagen-binding domain"
XX
XX WO200214505-A1.
XX
XX 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-JP007036.
XX
XX 15-AUG-2000; 2000JP-00246341.
XX
XX (TERU) TERUMO CORP.
XX
XX Ishikawa T, Kitajima T;
XX
XX WPI; 2002-257605/30.
XX
XX P-PSDB; ABB06747.
XX
XX Collagen-binding hybrid polypeptide, useful in drug delivery system for
XX functional polypeptides, formulating into complex with collagen to give
XX functionally modified collagen matrix as biomaterial for tissue
XX regeneration.
XX
XX Example 1; Page 59-61; 86pp; Japanese.
XX


```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-8 (1-20) x AX202071 (1-310)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 19 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGAAATTCCAATGGTGCCTTGTGCCAC 78

RESULT 2
AF312399 AF312399 615 bp mRNA linear PRI 07-NOV-2000
LOCUS Homo sapiens fibronectin mRNA, partial cds.
DEFINITION
ACCESSION AF312399
VERSION AF312399.1 GI:11119230
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 615)
AUTHORS Kornblith,A.R., Umezawa,K., Vibe-Pedersen,K. and Baralle,F.E.
TITLE Primary structure of human fibronectin: differential splicing may
Generate at least 10 polypeptides from a single gene
EMBO J. 4 (7), 1755-1759 (1985)
JOURNAL MEDLINE 85284965
PUBMED 2992939
REFERENCE 2 (bases 1 to 615)
AUTHORS Steffensen,B. and Martin,P.A.
TITLE Fibronectin collagen binding domain (modules I6, I11, I12, I7)
Unpublished
JOURNAL
REFERENCE 3 (bases 1 to 615)
AUTHORS Steffensen,B. and Martin,P.A.
TITLE Direct Submission
JOURNAL Submitted (10-OCT-2000) Periodontics, University of Texas Health
Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio,
TX 78229-3900, USA

FEATURES
source
1..615
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q34"
/cell_lines="gingival fibroblasts IMR90"
/note="identical to sequences found in gingival fibroblast
cell-lines WI38, MRC5, and Hs578T"
<1..>615

CDS
/note="spans modules I6, I11, I12, and I7; collagen
binding domain"
/codon_start=1
/product="fibronectin"
/protein_id="ANG30571.1"
/db_xref="GI:11119231"
/translation="GHCVTDSGVVYSGVMQLKQGNKQMLCTCLGNVGSQETAVTQ
TYGNSNGPCVLPFTYNGRTFYSCTEGRQDHLWCSTSNEDQDKYSECTDHTVL
VQTRGNGNGLCHPFLYNNHNYTDCSTSEGRDNNKWCSTTQNYDQKFGFCPMAA
HHEICTTNEGVMYRIGDQDQHDQHDGHWMRCTCVGNRGREWTCTA"
52
conflict
/citation=[1]
/replace="t"
149
conflict
/citation=[1]
/replace="t"
213
conflict
/citation=[1]
/replace="g"
317
conflict
/citation=[1]
/replace="a"

ORIGIN
```

```
Alignment Scores:
Pred. No.: 2.66e-08 Length: 615
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-581-651D-8 (1-20) x AF312399 (1-615)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 289 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGAAATTCCAATGGTGCCTTGTGCCAC 348

RESULT 3
BD142777
LOCUS BD142777 708 bp DNA linear PAT 18-SEP-2002
DEFINITION Collagen-binding hybrid polypeptide.
ACCESSION BD142777
VERSION BD142777.1 GI:23237722
KEYWORDS WO 0214505-A/3.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 708)
AUTHORS Ishikawa,T. and Kitajima,T.
TITLE Collagen-binding hybrid polypeptide
JOURNAL Patent: WO 0214505-A 3 21-FEB-2002;
TERUMO CORP.TETSUYA ISHIKAWA,TAKASHI KITAJIMA
OS Artificial Sequence
PN WO 0214505-A/3
PD 21-FEB-2002
PF 15-AUG-2001 WO 2001JP007036
PR 15-AUG-2000 JP 00P 246341
PI TETSUYA ISHIKAWA,TAKASHI KITAJIMA
PC C12N15/09,C07K14/78,C07K19/00,C12N1/21,C12N5/10,A61K38/00 CC
Description of Artificial Sequence: Human
Fibronectin Collagen-binding
CC Domain Location/Qualifiers
FH Key (16)..(693)
FT CDS (109)
FT conflict (206)
FT conflict (270)
FT conflict (374)
FT conflict (681)
FEATURES
source
1..708
/organism="synthetic construct"
/db_xref="taxon:32630"

ORIGIN

Alignment Scores:
Pred. No.: 3.02e-08 Length: 708
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-8 (1-20) x BD142777 (1-708)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 352 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGAAATTCCAATGGTGCCTTGTGCCAC 411

RESULT 4
CQ427970/c
LOCUS CQ427970 741 bp DNA linear PAT 28-JAN-2004
DEFINITION Sequence 13004 from Patent WO0151628.
ACCESSION CQ427970
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 7.06536 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTSDTTSNYEQDQ 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95.5	74.0	749	2 Q7Z391	Q7Z391 homo sapien
2	95.5	74.0	1103	2 Q6MZ4	Q6MZ4 homo sapien
3	95.5	74.0	2193	2 Q6MZ7	Q6MZ7 homo sapien
4	95.5	74.0	2240	2 Q6MDP8	Q6MDP8 homo sapien
5	95.5	74.0	2267	2 Q6MDP9	Q6MDP9 homo sapien
6	95.5	74.0	2296	2 Q6N0A6	Q6N0A6 homo sapien
7	95.5	74.0	2357	2 Q6MDT4	Q6MDT4 homo sapien
8	95.5	74.0	2386	1 F1NC HUMAN	Q6N025 homo sapien
9	95.5	74.0	2444	2 Q6N025	Q6N025 homo sapien
10	95.5	74.0	2477	1 F1NC MOUSE	P11276 mus musculus
11	95.5	74.0	2477	2 Q6MZ5	Q6MZ5 homo sapien
12	92.5	71.7	2265	1 F1NC BOVIN	P07589 bos taurus
13	88.5	68.6	2477	1 F1NC RAT	P04937 rattus norv
14	71.5	55.4	712	1 MM09 BOVIN	P52176 bos taurus
15	67.5	52.3	324	2 Q6N2B2	Q6N2B2 bos taurus
16	65.5	50.8	704	1 MM09 CANFA	Q18733 canis fami
17	64.5	50.0	208	2 Q6XSD1	Q9XSD1 canis fami
18	64.5	50.0	704	2 Q71U09	Q71U09 canis fami
19	64.5	50.0	2319	2 Q95MJ4	Q95MJ4 myotis luci
20	63.5	49.2	661	2 Q6GLE5	Q6GLE5 bos taurus
21	62.5	48.4	670	2 Q6DF16	Q6DF16 xenopus tro
22	62.5	48.4	586	2 Q6DE15	Q6DE15 gallus gall
23	62.5	48.4	2328	2 Q95MI9	Q95MI9 erinaceus e
24	61.5	47.7	671	2 Q6PF33	Q6PF33 xenopus lae
25	61.5	47.7	730	1 Q9W7L6	Q9W7L6 xenopus lae
26	61.5	47.7	730	1 MM09 MOUSE	P41245 mus musculu
27	61.5	47.7	730	2 Q80X18	Q80X18 mus musculu
28	61.5	47.7	2421	2 Q95MJ1	Q95MJ1 lemur catta
29	61.5	47.7	2481	1 F1NC_XENLA	Q91740 xenopus lae
30	61.5	47.7	2481	2 Q6GQAS	Q6GQAS xenopus lae
31	60.5	46.9	1463	1 PA2R_BOVIN	P49259 bos taurus

32	60.5	46.9	2420	2 Q95MI8	Q95MI8 cynocephalu
33	59.5	46.1	248	2 Q9BGL4	Q9BGL4 ovis aries
34	59.5	46.1	1326	2 Q13019	Q13019 homo sapien
35	59.5	46.1	1458	1 PA2R_RABIT	P49260 oryctolagus
36	59.5	46.1	1465	2 Q13018	Q13018 homo sapien
37	58.5	45.3	708	1 MM09_RAT	P50282 rattus norv
38	58.5	45.3	2437	2 Q95MJ3	Q95MJ3 oryctolagus
39	58.5	45.3	2473	2 Q95LC7	Q95LC7 tachyglossu
40	58.5	45.3	2487	2 Q9N1T0	Q9N1T0 ornithorhyn
41	58.5	45.3	2499	1 NPRI_BOVIN	P08169 bos taurus
42	57.5	44.6	2358	2 Q95MJ2	Q95MJ2 macropus ru
43	57.5	44.6	2425	2 Q95MJ0	Q95MJ0 tupia glis
44	56.5	43.8	385	2 Q9TUL8	Q9TUL8 equus cabal
45	56.5	43.8	632	2 Q9N1P6	Q9N1P6 canis fami

ALIGNMENTS

RESULT 1

Q7Z391 PRELIMINARY; PRT; 749 AA.
AC Q7Z391;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloecher H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -
DR HSSP; Q96KP7; 1PBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDP97F3ED2F0DE CRC64;

Query Match 74.0%; Score 95.5; DB 2; Length 749;

Best Local Similarity 55.3%; Pred. No. 3.1e-06;

Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPFTYNDRT-----DSTTSNYEQDQ 23

DB 451 PCVLPFTYNGRTFTYCTTEGRQDGHLMCSTTSNYEQDQ 488

RESULT 2

Q6MZ4 PRELIMINARY; PRT; 1103 AA.

AC Q6MZ4;

DT 05-JUL-2004 (TREMBLrel. 27, Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DE Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219)

DE (Fragment).

```
GN Name=DKFZp686K139; Synonym=DKFZp686P219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG The German Human cDNA Consortium;
RA Ansoerge W., Krieger S., Regliert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CAE46200.1; -.
DR EMBL; BX640802; CAE45885.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS0853; FN3; 4.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FECACAF634AD56 CRC64;

Query Match 74.0%; Score 95.5; DB 2; Length 1103;
Best Local Similarity 55.3%; Pred. No. 4.9e-06;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPFTYNDRT-----DSTTSNYEQDQ 23
DB 448 PCVLPFTYNGRTFYCTTGRQDGHLCWSTTSNYEQDQ 485

RESULT 3
Q6M2M7 PRELIMINARY; PRT; 2193 AA.
AC Q6M2M7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686O12165 (Fragment).
GN Name=DKFZp686O12165;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Human uterus endothel primary cell culture;
RG The German Human cDNA Consortium;
RA Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640999; CAE46002.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN_III.

GN Name=DKFZp686K139; Synonym=DKFZp686P219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Human cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749317; CAH18172.1; -.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 10.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
```

```
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 7.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 17.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 7.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 17.
DR PROSITE; PS00687; ALDEHYDE DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 7.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS0853; FN3; 17.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 2193 AA; 240641 MW; F876E93106540EF3 CRC64;

Query Match 74.0%; Score 95.5; DB 2; Length 2193;
Best Local Similarity 55.3%; Pred. No. 1.1e-05;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPFTYNDRT-----DSTTSNYEQDQ 23
DB 106 PCVLPFTYNGRTFYCTTGRQDGHLCWSTTSNYEQDQ 143

RESULT 4
Q68DP8 PRELIMINARY; PRT; 2240 AA.
AC Q68DP8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686H0342.
GN Name=DKFZp686H0342;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RG The German cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749317; CAH18172.1; -.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 10.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 1.74379 Seconds
(without alignments)
1269.066 Million cell updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95.5	74.0	2386	1 FNBU	fibronectin prec
2	92.5	71.7	2265	1 FNBO	fibronectin - bovi
3	88.5	68.6	2477	2 S14428	fibronectin prec
4	71.5	55.4	712	1 I46031	gelatinase B (EC 3
5	61.5	47.7	730	1 I52580	gelatinase B (EC 3
6	61.5	47.7	730	2 JCI456	gelatinase B (EC 3
7	61.5	47.7	2481	2 A43908	fibronectin - Afri
8	60.5	46.9	1463	2 A53210	phospholipase A2 r
9	59.5	46.1	1326	2 B56395	secretory phosphol
10	59.5	46.1	1458	1 A49707	phospholipase A2 r
11	59.5	46.1	1465	2 A56395	secretory phosphol
12	58.5	45.3	708	2 J43364	gelatinase B (EC 3
13	58.5	45.3	708	2 S62907	gelatinase B (EC 3
14	58.5	45.3	2499	1 A30788	mannose 6-phosphat
15	56.5	43.8	660	1 A28153	gelatinase A (EC 3
16	55.5	43.0	662	2 A42496	gelatinase A (EC 3
17	55.5	43.0	662	2 S34780	gelatinase A (EC 3
18	55.5	43.0	2491	1 A28372	insulin-like growt
19	54.5	42.2	603	2 S28941	coagulation factor
20	52.5	40.7	662	2 S70365	gelatinase A (EC 3
21	52.5	40.7	707	1 A53795	gelatinase B (EC 3
22	52.5	40.7	2482	2 I48922	cation-independent
23	52.5	40.7	2483	1 A49617	insulin-like growt
24	50.5	39.1	1487	2 S48719	phospholipase-A(2)
25	50	38.8	456	2 S45137	fushi tarazu segme
26	49.5	38.4	707	1 A34458	gelatinase B (EC 3
27	49	38.0	326	2 A47523	cartilage homeopro
28	49	38.0	2242	2 A57541	pyrimidine synthe
29	48.5	37.6	593	2 S45281	coagulation factor

RESULT 1

FNBU

fibronectin precursor [validated] - human
N/Alternate names: fibronectin splice form ED-A

C/Species: Homo sapiens (man)

C/Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004

C/Accession: A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22

R/Dean: D.C.; Bowlus, C.L.; Bourgeois, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A/Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A/Reference number: A26460; MUID:87175578; PMID:3031656

A/Accession: A26460

A/Molecule type: DNA

A/Residues: 1-49 <DEA>

A/Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:G182686; PIDN:AAA5337

R/Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A/Title: Evolution of the fibronectin gene.

A/Reference number: A26284; MUID:86111901; PMID:3003095

A/Accession: A26284

A/Molecule type: DNA

A/Residues: 1447-1540 <OLD>

A/Cross-references: GB:M12549; NID:G182688

A/Note: the authors translated the codon TTC for residue 1494 as Glu

R/Paoletti, G.; Henschliffe, C.; Sebastio, G.; Baralle, F.E.

Nucleic Acids Res. 15, 3545-3557, 1988

A/Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B

A/Reference number: S00848; MUID:88233940; PMID:3375063

A/Accession: S03917

A/Molecule type: DNA

A/Residues: 1594-1767, 'V', 1769-1783 <PAO>

A/Cross-references: EMBL:X07718; NID:G31402

A/Note: the authors translated the codon AAC for residue 1631 as Asp

R/Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.

FEBS Lett. 207, 287-291, 1986

A/Title: Donor and acceptor splice signals within an exon of the human fibronectin gene.

A/Reference number: A24854; MUID:87030929; PMID:3770201

A/Accession: A24854

A/Molecule type: DNA

A/Residues: 1992-2147 <VIB>

A/Cross-references: GB:X04530; NID:G31436

R/Gutman, A.; Yamada, K.M.; Kornblitt, A.

FEBS Lett. 207, 145-148, 1986

A/Title: Human fibronectin is synthesized as a pre-propolypeptide.

A/Reference number: A24476; MUID:87030890; PMID:3770189

A/Accession: A24476

A/Molecule type: mRNA

A/Residues: 1-14, 'Q', 16-38 <GUT>

R/Kornblitt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.

EMBO J. 4, 1753-1759, 1985

A/Title: Primary structure of human fibronectin: differential splicing may generate at

transcription regu
hepatocyte growth
probable dimethyl
hemocytin - silkwo
coagulation factor
mannose receptor,
hypothetical prote
C-8 sterol isomera
seminal fluid prot
flagellar M-ring p
F83508
hypothetical prote
AD3-like protein
hypothetical prote
zonadhesin - mouse

A:Reference number: A91008; MUID:85284965; PMID:2992939
A:Accession: A91008
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 32-1344, 1346-2080/2112-2386 <KOR>
A:Cross-references: GB:X02761
R:Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A:Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptide
A:Reference number: A93529; MUID:84272258; PMID:6462919
A:Accession: A93529
A:Molecule type: mRNA
A:Residues: 973-2080/2112-2386 <K02>
A:Cross-references: GB:X00739
R:Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A:Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell
A:Reference number: A21011; MUID:83290329; PMID:6688418
A:Accession: A21011
A:Molecule type: mRNA
A:Residues: 1434-1537 <OL2>
A:Cross-references: GB:K00055; NID:g182680; PIDN:AAA52459.1; PID:g182683
R:Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A:Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
A:Reference number: A90495; MUID:85280409; PMID:2992573
A:Accession: A90495
A:Molecule type: mRNA
A:Residues: 1594-2386 <BER>
A:Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697
R:Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A:Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A:Reference number: A22245; MUID:85231203; PMID:2989004
A:Accession: A22245
A:Molecule type: mRNA
A:Residues: 1948-2067 <UMB>
A:Cross-references: GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:g182706
A:Accession: B22245
A:Molecule type: mRNA
A:Residues: 1975-1991/2017-2039 <UM2>
A:Cross-references: GB:M27590
R:Sekiguchi, K.; Klos, A.M.; Kurechi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A:Title: Human liver fibronectin complementary DNAs: identification of two different me
A:Reference number: I52394; MUID:87026578; PMID:3021206
A:Accession: I65273
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1978-1990/2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
A:Cross-references: GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:g182704
R:Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A:Reference number: A21165; MUID:83221567; PMID:6304699
A:Accession: A21165
A:Molecule type: mRNA
A:Residues: 2291-2386 <K03>
A:Cross-references: GB:K00799; NID:g182681; PIDN:AAA52460.1; PID:g182684
R:Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A:Title: Primary structure of human plasma fibronectin.
A:Reference number: A92398; MUID:84032463; PMID:6630202
A:Accession: A92398
A:Molecule type: protein
A:Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>
R:Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A:Title: Further characterization of the binding of fibronectin to gelatin reveals the p
A:Reference number: S34791; MUID:93312001; PMID:8323285
A:Accession: S34791
A:Molecule type: protein
A:Residues: 291-300/551-560 <GAR2>

R:Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A:Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A:Reference number: A60904; MUID:87019725; PMID:3532418
A:Accession: A60904
A:Molecule type: protein
A:Residues: 293-301 <GR1>
R:Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A:Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl
A:Reference number: A23901; MUID:86008277; PMID:3900070
A:Accession: A23901
A:Molecule type: protein
A:Residues: 616-677, 'Q', 679-703, 'P', 'T' <CAL>
R:Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A:Title: The cell attachment domain of fibronectin. Determination of the primary structu
A:Reference number: A92386; MUID:82265604; PMID:7050098
A:Accession: A92386
A:Molecule type: protein
A:Residues: 1441-1548 <PIE>
A:Note: residues 1524-1527 are responsible for the cell-binding activity
R:Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A:Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A:Reference number: A32517; MUID:87241275; PMID:3593230
A:Accession: A32517
A:Molecule type: protein
A:Residues: 1589-1630, 'T', 1722-2058 <GAR3>
R:Trassel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A:Title: Human plasma fibronectin. Demonstration of structural differences between the A
A:Reference number: S14357; MUID:91190085; PMID:2012601
A:Accession: S14357
A:Molecule type: protein
A:Residues: 1614-1630, 'T', 1722-2081, 2113-2244 <TRE>
R:Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A:Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A:Reference number: A23891; MUID:85261459; PMID:4019516
A:Accession: A23891
A:Molecule type: protein
A:Residues: 2071-2080/2112-2356 <GAR4>
A:Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C:Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
ation, and transformation.
C:Genetics:
A:Gene: GDB:FNI
A:Cross-references: GDB:119135; OMIM:135600
A:Map position: 2q34-2q34
A:Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C:Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat
F:1-36/Domain: signal sequence #status predicted <SIG>
F:27-31/Domain: propeptide #status predicted <PRO>
F:32-2386/Product: fibronectin #status experimental
F:52-272/Domain: fibrin and heparin binding <FHB>
F:52-87/Domain: fibronectin type I repeat homology <1F1>
F:97-135/Domain: fibronectin type I repeat homology <1F2>
F:141-179/Domain: fibronectin type I repeat homology <1F3>
F:186-225/Domain: fibronectin type I repeat homology <1F4>
F:231-270/Domain: fibronectin type I repeat homology <1F5>
F:308-608/Domain: collagen binding <CBR>
F:308-342/Domain: fibronectin type I repeat homology <1F6>
F:360-401/Domain: fibronectin type II repeat homology <2F1>
F:420-461/Domain: fibronectin type II repeat homology <2F2>
F:470-508/Domain: fibronectin type I repeat homology <1F7>
F:518-555/Domain: fibronectin type I repeat homology <1F8>
F:561-599/Domain: fibronectin type I repeat homology <1F9>
F:609-692/Domain: fibronectin type III repeat homology <3FA>
F:616-706/Domain: heparin binding <HPB>
F:719-801/Domain: fibronectin type III repeat homology <3FB>
F:810-891/Domain: fibronectin type III repeat homology <3FC>

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 7.78693 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-7

Perfect score: 129

Sequence: 1 PCVLPFTYNDRTSTTSNYEQDQ 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	113	87.6	642	16	US-10-741-601-354
2	113	87.6	642	17	US-10-741-600-1066
3	95.5	74.0	228	15	US-10-344-634-4
4	95.5	74.0	285	15	US-10-344-634-16
5	95.5	74.0	343	9	US-09-934-706-1
6	95.5	74.0	386	15	US-10-344-634-14
7	95.5	74.0	400	9	US-09-934-706-5
8	95.5	74.0	463	15	US-10-144-194A-52
9	95.5	74.0	463	16	US-10-491-566-52
10	95.5	74.0	501	9	US-09-934-706-4
11	95.5	74.0	657	16	US-10-741-601-359

12	95.5	74.0	657	17	US-10-741-600-1072	Sequence 1072, Ap
13	95.5	74.0	984	16	US-10-741-601-356	Sequence 356, App
14	95.5	74.0	984	17	US-10-741-600-1069	Sequence 1069, Ap
15	95.5	74.0	1173	18	US-10-450-763-52634	Sequence 52634, A
16	95.5	74.0	2220	15	US-10-236-392-4	Sequence 4, Appli
17	95.5	74.0	2296	16	US-10-741-601-363	Sequence 363, App
18	95.5	74.0	2296	17	US-10-741-600-1075	Sequence 1075, Ap
19	95.5	74.0	2320	14	US-10-279-733-8	Sequence 8, Appli
20	95.5	74.0	2320	15	US-10-236-392-2	Sequence 2, Appli
21	95.5	74.0	2328	14	US-10-171-311-64	Sequence 64, Appl
22	95.5	74.0	2328	15	US-10-236-031B-70	Sequence 70, Appl
23	95.5	74.0	2328	15	US-10-374-979-98	Sequence 98, Appl
24	95.5	74.0	2328	15	US-10-182-336A-98	Sequence 98, Appl
25	95.5	74.0	2328	16	US-10-477-238A-677	Sequence 677, App
26	95.5	74.0	2328	16	US-10-680-287A-677	Sequence 677, App
27	95.5	74.0	2328	17	US-10-477-173-677	Sequence 677, App
28	95.5	74.0	2355	15	US-10-144-194A-104	Sequence 104, App
29	95.5	74.0	2355	15	US-10-360-101-235	Sequence 235, App
30	95.5	74.0	2355	15	US-10-447-161-3	Sequence 3, Appli
31	95.5	74.0	2355	16	US-10-734-564-94	Sequence 94, Appli
32	95.5	74.0	2355	16	US-10-741-601-357	Sequence 357, App
33	95.5	74.0	2355	16	US-10-741-601-366	Sequence 366, App
34	95.5	74.0	2355	16	US-10-491-566-104	Sequence 104, App
35	95.5	74.0	2355	17	US-10-741-600-1067	Sequence 1067, Ap
36	95.5	74.0	2355	17	US-10-741-600-1078	Sequence 1078, Ap
37	95.5	74.0	2355	17	US-10-852-335A-147	Sequence 147, App
38	95.5	74.0	2355	18	US-10-287-436A-436	Sequence 436, App
39	95.5	74.0	2355	18	US-10-287-436A-1137	Sequence 1137, Ap
40	95.5	74.0	2355	20	US-11-040-130-28	Sequence 28, Appli
41	95.5	74.0	2386	10	US-09-961-403-1	Sequence 1, Appli
42	95.5	74.0	2386	16	US-10-741-601-360	Sequence 360, App
43	95.5	74.0	2386	16	US-10-788-792-206	Sequence 206, App
44	95.5	74.0	2386	16	US-10-618-281-32	Sequence 32, Appl
45	95.5	74.0	2386	17	US-10-741-600-1071	Sequence 1071, Ap

ALIGNMENTS

RESULT 1
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match 87.6%; Score 113; DB 16; Length 642;
Best Local Similarity 91.3%; Pred. No. 4.1e-08;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PCVLPFTYNDRTSTTSNYEQDQ 23
DB 359 PCVLPFTYNGHTCSTTSNYEQDQ 381

RESULT 2
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

```
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match      87.6%; Score 113; DB 17; Length 642;
Best Local Similarity 91.3%; Pred. No. 4.1e-08;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PCVLPFTYNDRTDSTTSNYEQDQ 23
Db 359 PCVLPFTYNGRTCTTSNYEQDQ 381

RESULT 3
US-10-344-634-4
; Sequence 4, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Kitajima, Tetsuya
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human Fibronectin
; OTHER INFORMATION: Collagen-binding Domain
US-10-344-634-4

Query Match      74.0%; Score 95.5; DB 15; Length 228;
Best Local Similarity 55.3%; Pred. No. 5.3e-06;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPFTYNDRT-----DSTTSNYEQDQ 23
Db 70 PCVLPFTYNGRTFYCTTEGRQDGHLMCSTTSNYEQDQ 107

RESULT 4
US-10-344-634-16
; Sequence 16, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Kitajima, Tetsuya
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 16
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Collagen-binding
US-10-344-634-16

Query Match      74.0%; Score 95.5; DB 15; Length 285;
Best Local Similarity 55.3%; Pred. No. 6.8e-06;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPFTYNDRT-----DSTTSNYEQDQ 23
Db 70 PCVLPFTYNGRTFYCTTEGRQDGHLMCSTTSNYEQDQ 107

RESULT 5
US-09-934-706-1
; Sequence 1, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 1
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Modified Human
; OTHER INFORMATION: Fibronectin Collagen-Binding Domain
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN
; LOCATION: (2)...(341)
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: CONFLICT
; LOCATION: (69)
; NAME/KEY: CONFLICT
; LOCATION: (125)
US-09-934-706-1

Query Match      74.0%; Score 95.5; DB 9; Length 343;
Best Local Similarity 55.3%; Pred. No. 8.4e-06;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPFTYNDRT-----DSTTSNYEQDQ 23
Db 70 PCVLPFTYNGRTFYCTTEGRQDGHLMCSTTSNYEQDQ 107

RESULT 6
US-10-344-634-14
; Sequence 14, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Kitajima, Tetsuya
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 2.43529 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-7

Perfect score: 129

Sequence: 1 PCVLPTFTYNDRTSTTSNYEQDQ 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95.5	74.0	60	2	US-08-982-597A-18
2	95.5	74.0	60	3	US-09-136-218-26
3	95.5	74.0	2231	1	US-08-153-799-16
4	95.5	74.0	2324	1	US-08-283-857-1
5	95.5	74.0	2324	5	PCT-US95-09819-1
6	95.5	74.0	2327	6	5455158-1
7	95.5	74.0	2327	6	5455158-1
8	95.5	74.0	2386	2	US-09-016-366A-12
9	95.5	74.0	2386	4	US-09-961-403-1
10	95.5	74.0	2446	2	US-08-551-356-2
11	95.5	74.0	2446	5	PCT-US93-12687-2
12	60.5	46.9	1463	1	US-08-220-603A-11
13	59.5	45.1	1326	4	US-09-949-016-6806
14	59.5	46.1	1339	4	US-09-949-016-10448
15	59.5	46.1	1478	4	US-09-949-016-8315
16	56.5	43.8	429	4	US-09-194-468A-45
17	56.5	43.8	631	3	US-08-448-489-17
18	56.5	43.8	631	4	US-09-689-730-17
19	56.5	43.8	660	3	US-08-704-711A-18
20	56.5	43.8	660	3	US-09-521-220-18
21	56.5	43.8	660	3	US-09-391-104-19
22	56.5	43.8	660	4	US-09-917-254-89
23	56.5	43.8	660	4	US-09-949-016-6512
24	56.5	43.8	660	4	US-09-949-016-7937
25	55.5	43.0	2491	4	US-09-207-363-1
26	54	41.9	24	2	US-08-982-597A-23
27	54	41.9	24	3	US-09-136-218-23

28	52	40.3	16	2	US-08-982-597A-26	Sequence 26, Appl
29	52	40.3	16	3	US-09-136-218-26	Sequence 26, Appl
30	50.5	39.1	65	2	US-08-982-597A-19	Sequence 19, Appl
31	50.5	39.1	65	3	US-09-136-218-19	Sequence 19, Appl
32	50.5	39.1	1487	3	US-08-840-062-7	Sequence 7, Appl
33	49.5	38.4	707	3	US-08-704-711A-19	Sequence 19, Appl
34	49.5	38.4	707	3	US-09-521-220-19	Sequence 19, Appl
35	49.5	38.4	707	3	US-09-391-104-20	Sequence 20, Appl
36	49.5	38.4	707	4	US-09-949-016-6575	Sequence 6575, Ap
37	49.5	38.4	708	3	US-08-448-489-16	Sequence 16, Appl
38	49.5	38.4	708	4	US-09-689-730-16	Sequence 16, Appl
39	49.5	38.4	713	4	US-09-949-016-10629	Sequence 10629, A
40	49	38.0	373	4	US-08-248-796A-18953	Sequence 18953, A
41	48	37.2	410	1	US-08-091-519-13	Sequence 13, Appl
42	48	37.2	410	1	US-08-442-043A-13	Sequence 13, Appl
43	48	37.2	410	3	US-09-173-151A-27	Sequence 27, Appl
44	48	37.2	410	4	US-09-461-908-13	Sequence 13, Appl
45	48	37.2	410	4	US-08-441-893A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-982-597A-18
; Sequence 18, Application US/08982597A
; Patent No. 5932693
; GENERAL INFORMATION:
; APPLICANT: Santoro, Samuel A.
; APPLICANT: Staats, William D.
; TITLE OF INVENTION: Antithrombotic Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer
; STREET: 800 No. 5932693th Lindbergh Blvd.
; CITY: St. Louis
; STATE: MO
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,597A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,542
; FILING DATE: 10-DEC-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WJ-3002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-694-3117
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-982-597A-18

Query Match 74.0%; Score 95.5; DB 2; Length 60;
Best Local Similarity 55.3%; Pred. No. 1.7e-07;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 PCVLPTFTYNDRT-----DSSTTSNYEQDQ 23
| | | | | | | | | | | | | | | | | | | | |
Db 15 PCVLPTFTYNGRTFYSCCTTEGQDGLMWCSTTSNYEQDQ 52

RESULT 2
US-09-136-218-18
; Sequence 18, Application US/09136218
; Patent No. 6083914
; GENERAL INFORMATION:
; APPLICANT: Santoro, Samuel A.
; APPLICANT: Staatz, William D.
; TITLE OF INVENTION: Antithrombotic Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer
; STREET: 800 No. 6083914th Lindbergh Blvd.
; CITY: St. Louis
; STATE: MO
; COUNTRY: USA
; ZIP: 63167

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09136,218
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,597

; FILING DATE:
; APPLICATION NUMBER: 60/032,542
; FILING DATE: 10-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-3002
; TELEPHONE: 314-694-3117
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-136-218-18

Query Match 74.0%; Score 95.5; DB 3; Length 60;
Best Local Similarity 55.3%; Pred. No. 1.7e-07;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 PCVLPTNYNDRT-----DSTTSNYEQDQ 23
Db 15 PCVLPTNYGRFYSCCTTEGRQDGHLMWCSTTSNYEQDQ 52

RESULT 3
US-08-153-799-16
; Sequence 16, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2231 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Protein
LOCATION: 1..2231
OTHER INFORMATION: /note= "Human fibronectin"
US-08-153-799-16

Query Match 74.0%; Score 95.5; DB 1; Length 2231;
Best Local Similarity 55.3%; Pred. No. 1e-05;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 PCVLPTNYNDRT-----DSTTSNYEQDQ 23
Db 328 PCVLPTNYGRFYSCCTTEGRQDGHLMWCSTTSNYEQDQ 365

RESULT 4
US-08-283-857-1
; Sequence 1, Application US/08283857
; Patent No. 5752742
; GENERAL INFORMATION:
; APPLICANT: GOLD, Leslie I.
; APPLICANT: ROSTAGNO, Agueda A.
; APPLICANT: BARON, Martin
; APPLICANT: CAMPBELL, Iain D.
; APPLICANT: WILLIAMS, Michael, J.
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 8.5085 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651D-7

Perfect score: 129

Sequence: 1 PCVLPFTYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_16Dec04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	129	100.0	23	2	AAY28912
2	129	100.0	24	2	AAY28904
3	129	100.0	660	2	AAY28901
4	122	94.6	642	8	ADR67316
5	122	94.6	642	8	ADS17489
6	122	94.6	642	8	ADR97658
7	113	87.6	642	8	ADQ39403
8	95.5	74.0	39	2	AAY28908
9	95.5	74.0	60	2	AAW62371
10	95.5	74.0	228	5	ABB06747
11	95.5	74.0	285	5	ABB06751
12	95.5	74.0	343	3	ABB08505
13	95.5	74.0	343	3	ABB07961
14	95.5	74.0	386	5	ABB06750
15	95.5	74.0	400	3	ABB08509
16	95.5	74.0	461	5	ABB07964
17	95.5	74.0	463	6	ABR58303
18	95.5	74.0	473	5	ABB07965
19	95.5	74.0	501	3	ABB08508
20	95.5	74.0	657	8	ADQ39409
21	95.5	74.0	720	2	AAY28914
22	95.5	74.0	951	8	ADS19044
23	95.5	74.0	984	8	ADQ39406
24	95.5	74.0	1173	4	ABG22275
25	95.5	74.0	1179	8	ADP75952

ALIGNMENTS

RESULT 1

AAY28912
ID AAY28912 standard; peptide; 23 AA.

XX AC AAY28912;

XX DT 21-SBP-1999 (first entry)

XX DE MSF 1-alpha peptide epitope.

XX KW Migration stimulatory factor; MSF; cell migration; modulation; human;
wound healing; scarring; MSF1-alpha; epitope; fibronectin.

XX OS Synthetic.

XX OS Homo sapiens.

XX FN WO9931233-A1.

XX PD 24-JUN-1999.

XX PF 15-DEC-1998; 98WO-GB003766.

XX PR 16-DEC-1997; 97GB-00026539.

XX PA (UYDU-) UNIV DUNDEE.

XX PI Schor SL, Schor AM;

XX DR WPI; 1999-430039/36.

XX PT Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX PS Claim 29; Page 61; 86pp; English.

XX CC The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. Sequences AAY28912-913 CC represent peptide epitopes of MSF against which monoclonal antibodies can be raised

XX SQ Sequence 23 AA;

Query Match 100.0%; Score 129; DB 2; Length 23;

Best Local Similarity 100.0%; Pred. No. 1.3e-12;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Migration stimulatory factor; MSP; cell migration; modulation; human; wound healing; scarring; MSF1-alpha.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 478.34 Seconds
(without alignment)
1830.243 Million cell updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092223_28789/app_query.fasta_1.1834
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651.qcgn 1 1 8180 @runat_07112005_092223_28789 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_hc:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gse1:
9: gb_gse2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95.5	74.0	380	CN701692	CN701692 E0452H02-
2	95.5	74.0	385	AL589418	AL589418 DKF2p451P
3	95.5	74.0	392	BE079849	BE079849 RC6-BT062
4	95.5	74.0	400	BE122858	BE122858 14_08 Hum
5	95.5	74.0	411	BE642740	BE642740 DKF2p781G
6	95.5	74.0	431	BE815722	BE815722 PM3-BN017
7	95.5	74.0	431	CD550770	CD550770 B0318C06-
8	95.5	74.0	458	AJ661698	AJ661698 AJ661698
9	95.5	74.0	471	BG007031	BG007031 RC1-GN023

10	95.5	74.0	471	5	BX495442
c 11	95.5	74.0	475	5	BX007376
12	95.5	74.0	478	5	BX495600
13	95.5	74.0	484	4	BF986449
14	95.5	74.0	505	7	CF166960
c 15	95.5	74.0	507	4	BF985411
16	95.5	74.0	510	2	AW368063
17	95.5	74.0	513	5	BX642768
18	95.5	74.0	521	1	AJ668111
c 19	95.5	74.0	528	7	CF793207
20	95.5	74.0	538	1	AJ685866
21	95.5	74.0	552	4	BG945246
22	95.5	74.0	558	7	CN419467
c 23	95.5	74.0	564	2	BF229278
24	95.5	74.0	568	5	BQ340570
25	95.5	74.0	605	7	CN481619
26	95.5	74.0	618	7	CN482377
27	95.5	74.0	620	7	CN481412
28	95.5	74.0	622	7	CN419625
29	95.5	74.0	626	7	CF907680
30	95.5	74.0	635	5	BQ312550
c 31	95.5	74.0	645	4	BG945249
32	95.5	74.0	645	7	CN419499
33	95.5	74.0	646	2	AW321646
34	95.5	74.0	655	7	CN419546
35	95.5	74.0	657	7	CN419503
36	95.5	74.0	702	7	CN419472
37	95.5	74.0	705	1	AU140899
c 38	95.5	74.0	725	5	BQ574857
39	95.5	74.0	748	1	AU140948
40	95.5	74.0	748	1	AU140991
41	95.5	74.0	758	1	AU140556
42	95.5	74.0	759	1	AU140735
c 43	95.5	74.0	780	7	CN162843
44	95.5	74.0	781	7	CN161442
45	95.5	74.0	786	4	BI149026

ALIGNMENTS

RESULT 1

CN701692

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

CN701692 E0452H02-5 NIA Mouse E11.5 whole embryo cDNA library (Long) Mus musculus cDNA clone NIA:E0452H02 IMAGE:30873589 5', mRNA sequence.

CN701692.1 GI:47470441

EST.

Mus musculus (house mouse)

1 (bases 1 to 380)

Sharov, A., Piao, Y., Matoba, R., Dudekula, D. B., Qian, Y., VanBuren, V., Falco, G., Martin, P. R., Stagg, C. A., Basse, U. C., Wang, X., Carter, M. G., Hamatani, T., Alba, K., Akutsu, H., Sharova, L., Tanaka, T. S., Kimber, W. L., Yoshikawa, T., Jaradat, S. A., Pantano, S., Nagaraja, R., Boheler, K. R., Taub, D., Hodes, R. J., Longo, D. L., Schlesinger, D., Keller, J., Klotz, E., Kelsae, G., Umezawa, A., Versco, A. L., Rossant, J., Kunath, T., Hogan, B. L., Curci, A., D'Urso, M., Kelsae, J., Hide, W. and Ko, M. S.

TITLE

JOURNAL

COMMENT

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igun.grc.nia.nih.gov

Plate: E0452 row: H column: 02

Seq primer: M13 Reverse

High quality sequence stop: 380

POLYA-No. Location/Qualifiers

source

```
1. .380
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="niaEST:E0452H02-5"
/db_xref="taxon:10090"
/clone="NIA:E0452H02 IMAGE:30873589"
/tissue_type="whole embryo including extraembryonic
tissues at 11.5-days postcoitum"
/dev_stage="E11.5"
/lab_host="DH10B"
/clone_lib="NIA Mouse E11.5 whole embryo cDNA library
(long)"
/notes="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were extracted from a pool of 3 embryos at 11.5-days
postcoitum. Double-stranded cDNAs were synthesized with an
Oligo(dT) primer [Invitrogen]:
5'-PCACTAGTCTAGATCGGAGCGCGCCCTTTT-3' from
2ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lene-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 3.3Kb. The library was
constructed by Yulan Piao."
```

ORIGIN

Alignment Scores:

```
Pred. No.: 7,91e-06 Length: 380
Score: 95.50 Matches: 21
Percent Similarity: 55.26% Conservative: 0
Best Local Similarity: 55.26% Mismatches: 2
Query Match: 74.03% Indels: 15
DB: 7 Gaps: 1
```

US-09-581-651D-7 (1-23) x CN701692 (1-380)

```
Qy 1 ProCysValLeuProPheThrTyRAsnAspArgThr----- 12
Db 219 CCCTGTGCTCTCCGTCACCTACCAACGGTAGGACCTTCTATTCTGCACCAACGAGGG 278
Qy 13 -----AspSerThrThrSerAsnTyRGlulnAspGln 23
Db 279 CGGCAAGACGGACATCTGTGGTGTAGCACCACTTCATTACCAACAGACCCAG 332
```

RESULT 2

```
AL589418 385 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp451P1415 x1 451 (synonym: hlcc1) spinal cord Homo sapiens cDNA
DEFINITION clone DKFZp451P1415 5', mRNA sequence.
ACCESSION AL589418
VERSION AL589418.1 GI:13243190
KEYWORDS EST.
```

SOURCE

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 385)
AUTHORS Bloecker, H., Boecker, M., Brandt, P., Mewes, W., Weil, B. and
Wiemann, S.
TITLE EST (Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Weil, B. and
Wiemann, S.)

JOURNAL
COMMENT

Unpublished (1999)
Contact: MIPS

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.

No sl sequence available.

This clone (DKFZp451P1415) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

```
1. .385
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp451P1415"
/tissue_type="human spinal cord"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="451 (synonym: hlcc1) spinal cord"
/notes="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 8.03e-06 Length: 385
Score: 95.50 Matches: 21
Percent Similarity: 55.26% Conservative: 0
Best Local Similarity: 55.26% Mismatches: 2
Query Match: 74.03% Indels: 15
DB: 1 Gaps: 1
```

US-09-581-651D-7 (1-23) x AL589418 (1-385)

```
Qy 1 ProCysValLeuProPheThrTyRAsnAspArgThr----- 12
Db 133 CCATGTGTCTTACCATTCACCTACCAATGGCAGCGTTCTACTCTGCACCAAGAGGG 192
Qy 13 -----AspSerThrThrSerAsnTyRGlulnAspGln 23
Db 193 CGACAGGACGGACATCTTTGGTGCAGCACCACTTCGAATTATGACGAGCCAG 246
```

RESULT 3

LOCUS

BE079849 392 bp mRNA linear EST 12-JUN-2000
RC6-BT0627-220300-012-F02 BT0627 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BE079849 GI:8470133

VERSION

BE079849.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 392)

REFERENCE

AUTHORS

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 118.097 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTDTSNRYEQDQ 23

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1834
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALICN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPSTZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09581651@cgn1.1.1041 @runat_07112005_092226_28867
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	87.6	2127	17	US-10-210-120-49
2	113	87.6	2127	22	US-10-956-157-4298
3	113	87.6	2127	24	US-10-909-035-49
4	113	87.6	2443	20	US-10-741-601-70
5	113	87.6	2443	22	US-10-741-600-238
6	95.5	74.0	504	9	US-09-815-343-273
7	95.5	74.0	504	19	US-10-097-105-273
8	95.5	74.0	708	19	US-10-344-634-3
9	95.5	74.0	731	15	US-10-198-846-8918
10	95.5	74.0	879	19	US-10-344-634-15
11	95.5	74.0	1053	9	US-09-934-706-8
12	95.5	74.0	1182	19	US-10-344-634-13
13	95.5	74.0	1189	15	US-10-198-846-11556
14	95.5	74.0	1224	9	US-09-934-706-16
15	95.5	74.0	1527	9	US-09-934-706-15
16	95.5	74.0	2488	20	US-10-741-601-75
17	95.5	74.0	2488	22	US-10-741-600-244
18	95.5	74.0	3522	24	US-10-450-763-22266
19	95.5	74.0	4295	18	US-10-144-1948-51
20	95.5	74.0	4295	21	US-10-491-566-51
21	95.5	74.0	6510	20	US-10-741-601-72
22	95.5	74.0	6510	22	US-10-741-600-241
23	95.5	74.0	6988	19	US-10-236-392-1
24	95.5	74.0	7361	19	US-10-236-392-3
25	95.5	74.0	7677	22	US-10-956-157-4995
26	95.5	74.0	7679	22	US-10-831-704-38
27	95.5	74.0	7680	9	US-09-964-824A-574
28	95.5	74.0	7680	15	US-10-171-311-63
29	95.5	74.0	7680	18	US-10-236-031B-69
30	95.5	74.0	7680	19	US-10-374-979-75
31	95.5	74.0	7680	18	US-10-182-936A-75
32	95.5	74.0	7680	19	US-10-641-643-1289
33	95.5	74.0	7680	20	US-10-717-597-222
34	95.5	74.0	7680	20	US-10-788-792-79
35	95.5	74.0	7680	21	US-10-477-238A-654
36	95.5	74.0	7680	21	US-10-680-287A-654
37	95.5	74.0	7680	22	US-10-278-698-88
38	95.5	74.0	7680	22	US-10-478-698-603
39	95.5	74.0	7680	22	US-10-843-641A-5877
40	95.5	74.0	7680	22	US-10-477-173-654
41	95.5	74.0	7680	22	US-10-852-335A-52
42	95.5	74.0	7705	18	US-10-447-161-4
43	95.5	74.0	7795	16	US-10-084-817-2
44	95.5	74.0	7823	20	US-10-741-601-77
45	95.5	74.0	7823	22	US-10-741-600-245

ALIGNMENTS

RESULT 1
US-10-210-120-49
; Sequence 49, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:
; APPLICANT: Chinaiyan, Arul M.
; APPLICANT: Rubin, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UN-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02

```
/ PRIOR APPLICATION NUMBER: US 60/334,468
/ PRIOR FILING DATE: 2001-11-15
/ NUMBER OF SEQ ID NOS: 123
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 49
/ LENGTH: 2127
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-210-120-49

Alignment Scores:
Pred. No.: 5,23e-09 Length: 2127
Score: 113.00 Matches: 21
Percent Similarity: 91.30% Conservative: 0
Best Local Similarity: 91.30% Mismatches: 2
Query Match: 87.60% Indels: 0
DB: 17 Gaps: 0

US-09-581-651D-7 (1-23) x US-10-210-120-49 (1-2127)

Qy 1 ProCysValLeuProPheThrTyraAspArgThrAspSerThrSerAsnTyrGlu 20
Db 1131 CCATGTGCTTACCATTACCTACATGGCAGGACGTGCAGCAACCTTCGAATTATGAG 1190

Qy 21 GlnAspGln 23
Db 1191 CAGGACCAG 1199

RESULT 2
US-10-956-157-4288
/ Sequence 4288, Application US/10956157
/ Publication No. US20050118625A1
/ GENERAL INFORMATION:
/ APPLICANT: Wyeth
/ APPLICANT: Mounts, William
/ TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
/ FILE REFERENCE: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
/ FILE REFERENCE: 031896-043000 (AM 101081)
/ CURRENT APPLICATION NUMBER: US/10/956,157
/ CURRENT FILING DATE: 2004-10-04
/ NUMBER OF SEQ ID NOS: 319805
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 4288
/ LENGTH: 2127
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-956-157-4288

Alignment Scores:
Pred. No.: 5,23e-09 Length: 2127
Score: 113.00 Matches: 21
Percent Similarity: 91.30% Conservative: 0
Best Local Similarity: 91.30% Mismatches: 2
Query Match: 87.60% Indels: 0
DB: 22 Gaps: 0

US-09-581-651D-7 (1-23) x US-10-956-157-4288 (1-2127)

Qy 1 ProCysValLeuProPheThrTyraAspArgThrAspSerThrSerAsnTyrGlu 20
Db 1131 CCATGTGCTTACCATTACCTACATGGCAGGACGTGCAGCAACCTTCGAATTATGAG 1190

Qy 21 GlnAspGln 23
Db 1191 CAGGACCAG 1199

RESULT 3
US-10-909-035-49
/ Sequence 49, Application US/10909035
/ Publication No. US20050136493A1
/ GENERAL INFORMATION:
/ APPLICANT: Rubin, Mark A.
/ APPLICANT: Chinnaiyan, Arul M.

/ APPLICANT: Laxman, Bharathi
/ APPLICANT: Sreekumar, Arun
/ TITLE OF INVENTION: AMACR Cancer Markers
/ FILE REFERENCE: UM-09098
/ CURRENT APPLICATION NUMBER: US/10/909,035
/ CURRENT FILING DATE: 2004-07-30
/ NUMBER OF SEQ ID NOS: 123
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 49
/ LENGTH: 2127
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-909-035-49

Alignment Scores:
Pred. No.: 5,23e-09 Length: 2127
Score: 113.00 Matches: 21
Percent Similarity: 91.30% Conservative: 0
Best Local Similarity: 91.30% Mismatches: 2
Query Match: 87.60% Indels: 0
DB: 24 Gaps: 0

US-09-581-651D-7 (1-23) x US-10-909-035-49 (1-2127)

Qy 1 ProCysValLeuProPheThrTyraAspArgThrAspSerThrSerAsnTyrGlu 20
Db 1131 CCATGTGCTTACCATTACCTACATGGCAGGACGTGCAGCAACCTTCGAATTATGAG 1190

Qy 21 GlnAspGln 23
Db 1191 CAGGACCAG 1199

RESULT 4
US-10-741-601-70
/ Sequence 70, Application US/10741601
/ Publication No. US20040168519A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001500
/ CURRENT APPLICATION NUMBER: US/10/741,601
/ CURRENT FILING DATE: 2003-12-22
/ NUMBER OF SEQ ID NOS: 26415
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 70
/ LENGTH: 2443
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-741-601-70

Alignment Scores:
Pred. No.: 6,21e-09 Length: 2443
Score: 113.00 Matches: 21
Percent Similarity: 91.30% Conservative: 0
Best Local Similarity: 91.30% Mismatches: 2
Query Match: 87.60% Indels: 0
DB: 20 Gaps: 0

US-09-581-651D-7 (1-23) x US-10-741-601-70 (1-2443)

Qy 1 ProCysValLeuProPheThrTyraAspArgThrAspSerThrSerAsnTyrGlu 20
Db 1447 CCATGTGCTTACCATTACCTACATGGCAGGACGTGCAGCAACCTTCGAATTATGAG 1506

Qy 21 GlnAspGln 23
Db 1507 CAGGACCAG 1515

RESULT 5
US-10-741-600-238
/ Sequence 238, Application US/10741600
/ Publication No. US20050026169A1
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:49:16 ; Search time 23.1503 Seconds
(without alignments)
1625.652 Million cell updates/sec

Title: US-09-581-651D-7

Perfect score: 129

Sequence: 1 PCVLPFTYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPIO-spool_h/US09581651/runat_07112005_092224_28801/app_query.fasta_1.1834
-DB=Issued Patents NA -QWMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCD=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651 -CGEN 1 1 187 @runat_07112005_092224_28801 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95.5	74.0	7679	US-09-220-132-38	Sequence 38, Appl
2	95.5	74.0	7680	US-09-023-655-1289	Sequence 1289, Ap
3	95.5	74.0	7680	PCT-US95-09819-6	Sequence 6, Appli
4	95.5	74.0	7705	US-08-259-589-16	Sequence 16, Appl
5	95.5	74.0	7705	US-08-826-885-16	Sequence 16, Appl
6	95.5	74.0	7705	US-08-259-589-16	Sequence 16, Appl
7	95.5	74.0	7705	US-08-259-589-16	Sequence 16, Appl
8	95.5	74.0	7803	US-08-551-356-1	Sequence 1, Appli
9	95.5	74.0	7803	PCT-US93-12887-1	Sequence 1, Appli
10	95.5	74.0	8044	US-09-566-921-135	Sequence 135, App
11	88.5	68.6	986	US-07-637-250A-8	Sequence 8, Appli
12	88.5	68.6	986	US-08-145-061-8	Sequence 8, Appli

13	60.5	46.9	4978	1	US-08-220-603A-1	Sequence 1, Appli
c 14	59.5	46.1	601	4	US-09-949-016-36296	Sequence 36296, A
c 15	59.5	46.1	601	4	US-09-949-016-86573	Sequence 86573, A
c 16	59.5	46.1	601	4	US-09-949-016-163266	Sequence 163266,
17	59.5	46.1	4621	4	US-09-949-016-4577	Sequence 4577, Ap
18	59.5	46.1	4627	4	US-09-949-016-935	Sequence 935, App
19	59.5	46.1	5627	4	US-09-949-016-2444	Sequence 2444, Ap
20	59.5	46.1	5633	4	US-09-023-655-1490	Sequence 1490, Ap
21	59.5	46.1	119930	4	US-09-949-016-12677	Sequence 12677, A
22	59.5	46.1	119931	4	US-09-949-016-16319	Sequence 16319, A
23	59.5	46.1	125536	4	US-09-949-016-14186	Sequence 14186, A
24	56.5	43.8	601	4	US-09-949-016-30321	Sequence 30321, A
25	56.5	43.8	601	4	US-09-949-016-70165	Sequence 70165, A
26	56.5	43.8	1983	4	US-09-949-016-2066	Sequence 2066, App
27	56.5	43.8	3069	4	US-09-949-016-641	Sequence 641, App
28	56.5	43.8	29954	4	US-09-949-016-13808	Sequence 13808, A
29	56.5	43.8	31040	4	US-09-949-016-12383	Sequence 12383, A
30	52	40.3	1383	4	US-08-248-796A-2647	Sequence 2647, Ap
c 31	50	38.8	601	4	US-09-949-016-42379	Sequence 42379, A
c 32	50	38.8	601	4	US-09-949-016-42380	Sequence 42380, A
33	50	38.8	601	4	US-09-949-016-96535	Sequence 96535, A
34	50	38.8	601	4	US-09-949-016-96801	Sequence 96801, A
35	50	38.8	601	4	US-09-949-016-97067	Sequence 97067, A
36	50	38.8	601	4	US-09-949-016-97333	Sequence 97333, A
37	50	38.8	601	4	US-09-949-016-97599	Sequence 97599, A
38	50	38.8	601	4	US-09-949-016-97865	Sequence 97865, A
39	50	38.8	601	4	US-09-949-016-98131	Sequence 98131, A
40	50	38.8	601	4	US-09-949-016-98397	Sequence 98397, A
41	50	38.8	601	4	US-09-949-016-98663	Sequence 98663, A
42	50	38.8	601	4	US-09-949-016-98929	Sequence 98929, A
43	50	38.8	601	4	US-09-949-016-99195	Sequence 99195, A
44	50	38.8	601	4	US-09-949-016-99461	Sequence 99461, A
45	50	38.8	601	4	US-09-949-016-99727	Sequence 99727, A

ALIGNMENTS

RESULT 1

US-09-220-132-38
; Sequence 38, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-38

Alignment Scores:
Pred. No.: 9.42e-07 Length: 7679
Score: 95.50 Matches: 21
Percent Similarity: 55.26% Conservative: 0
Best Local Similarity: 55.26% Mismatches: 2
Query Match: 74.03% Indels: 15
DB: 4 Gaps: 1

US-09-581-651D-7 (1-23) x US-09-220-132-38 (1-7679)

Qy 1 ProCysValLeuProPheThrTyrAsnAspArgThr-----12

Db 995 CCATGTGTCTTACCATTCACCTACAAATGGCAGACGTTCTTACTCTGCACCCAGGAGGG 1054

Qy 13 -----AspSerThrThrSerAsnTyrGluGlnAspGln 23
Db 1055 CGACGAGCGGACATCTTTGGTGCAGCACTTCGAATTATGACGAGGACCAG 1108

RESULT 2
US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: G31396
; LIBRARY: GENBANK
; US-09-023-655-1289

Alignment Scores:
Pred. No.: 9.42e-07 Length: 7680
Score: 95.50 Matches: 21
Percent Similarity: 55.26% Conservative: 0
Best Local Similarity: 55.26% Mismatches: 2
Query Match: 74.03% Indels: 15
Gaps: 1

US-09-581-651D-7 (1-23) x US-09-023-655-1289 (1-7680)

Qy 1 ProCysValLeuProPheThrTyrAsnAspArgThr----- 12
Db 995 CCATGTGCTTACCATTCACCTACATGCGAGGAGTTCTACTCTCGACACCGAAGG 1054

Qy 13 -----AspSerThrThrSerAsnTyrGluGlnAspGln 23
Db 1055 CGACGAGCGGACATCTTTGGTGCAGCACTTCGAATTATGACGAGGACCAG 1108

RESULT 3

PCT-US95-09819-6
; Sequence 6, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US95-09819-6

Alignment Scores:
Pred. No.: 9.42e-07 Length: 7680
Score: 95.50 Matches: 21
Percent Similarity: 55.26% Conservative: 0
Best Local Similarity: 55.26% Mismatches: 2
Query Match: 74.03% Indels: 15
Gaps: 1

US-09-581-651D-7 (1-23) x PCT-US95-09819-6 (1-7680)

Qy 1 ProCysValLeuProPheThrTyrAsnAspArgThr----- 12
Db 995 CCATGTGCTTACCATTCACCTACATGCGAGGAGTTCTACTCTCGACACCGAAGG 1054

Qy 13 -----AspSerThrThrSerAsnTyrGluGlnAspGln 23
Db 1055 CGACGAGCGGACATCTTTGGTGCAGCACTTCGAATTATGACGAGGACCAG 1108

RESULT 4
US-08-259-569-16
; Sequence 16, Application US/08259569
; Patent No. 5679320
; GENERAL INFORMATION:
; APPLICANT: Vogel, Tikva
; APPLICANT: Levanon, Avigdor
; APPLICANT: Werber, Moshe
; APPLICANT: Guy, Rachel
; APPLICANT: Panet, Amos
; APPLICANT: Hartman, Jacob
; APPLICANT: Shaked, Hadassa
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:30:56 ; Search time 75.4039 Seconds
(without alignments)
1805.663 Million cell updates/sec

Title: US-09-581-651D-7

Perfect score: 129

Sequence: 1 PCVLPFTYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/Cgn2.1/USPTO.spool_h/US09581651/runat_07112005_092222_28772/app_query.fasta.1.1834
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651 @CGN 1.1.1052 @runat_07112005_092222_28772 -NCPUP=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04:.*
1: geneseqn1980e:.*
2: geneseqn1990e:.*
3: geneseqn2000e:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	2147	2 AAX81299	Aax81299 Human mig
2	122	94.6	1929	13 ADS17488	Adsl17488 Nucleotid
3	122	94.6	1929	13 ADR97657	Adr97657 Human fib
4	122	94.6	2358	13 ADR67201	Adr67201 Human bla
5	113	87.6	2127	10 ADD18477	Add18477 Human pro

6	113	87.6	2443	13	ADQ38575	Adq38575 Human SNP
7	95.5	74.0	504	4	AAS57597	Aas57597 cDNA #273
8	95.5	74.0	706	13	ADS19041	Adsl19041 Human fib
9	95.5	74.0	708	6	ABL50261	Ab150261 Human fib
10	95.5	74.0	731	11	ACN87768	Acn87768 Breast ca
11	95.5	74.0	741	4	ABL20536	Ab120536 Human bre
12	95.5	74.0	879	6	ABL50269	Ab150269 Collagen-
13	95.5	74.0	1051	13	ADS19042	Adsl19042 Human fib
14	95.5	74.0	1053	3	AAA64263	Aaa64263 cDNA enco
15	95.5	74.0	1053	6	ABL41020	Ab141020 Modified
16	95.5	74.0	1182	6	ABL50268	Ab150268 Collagen-
17	95.5	74.0	1189	11	ACN90406	Acn90406 Breast ca
18	95.5	74.0	1224	3	AAA64271	Aaa64271 Fibronect
19	95.5	74.0	1407	6	ABL41027	Ab141027 Human fib
20	95.5	74.0	1443	6	ABL41028	Ab141028 Human fib
21	95.5	74.0	1527	3	AAA64270	Aaa64270 Fibronect
22	95.5	74.0	2488	13	ADQ38581	Adq38581 Human SNP
23	95.5	74.0	2856	13	ADS19043	Adsl19043 Chimeric
24	95.5	74.0	3522	5	AAS86462	Aas86462 DNA encod
25	95.5	74.0	3540	12	ADP75954	Adp75954 Human min
26	95.5	74.0	3669	12	ADP29134	Adp29134 Human sec
27	95.5	74.0	4080	12	ADP75958	Adp75958 Human leu
28	95.5	74.0	4295	8	ACC72037	Acc72037 BCU0770 g
29	95.5	74.0	4860	3	AAA35009	Aaa35009 Human ade
30	95.5	74.0	6510	13	ADQ38578	Adq38578 Human SNP
31	95.5	74.0	6816	10	ABX74443	Abx74443 Human CDN
32	95.5	74.0	6988	9	ACD06169	Acd06169 Human CDN
33	95.5	74.0	7020	13	ADR90518	Adr90518 Human CDN
34	95.5	74.0	7049	12	ADN95947	Adn95947 Human NOV
35	95.5	74.0	7242	13	ADR66637	Adr66637 Human pro
36	95.5	74.0	7242	13	ADR65953	Adr65953 Human pro
37	95.5	74.0	7361	9	ACD06170	Acd06170 Human CDN
38	95.5	74.0	7550	8	ACC00412	Acc00412 Human cel
39	95.5	74.0	7677	13	ACN37820	Acn37820 Tumour-as
40	95.5	74.0	7679	10	ADB31322	Adb31322 Testoster
41	95.5	74.0	7680	2	AAT17551	Aat17551 Human fib
42	95.5	74.0	7680	3	AAF21131	Aaf21131 Human low
43	95.5	74.0	7680	5	ABA82689	Abas2689 Fibronect
44	95.5	74.0	7680	6	ABL67540	Ab167540 Thyroid c
45	95.5	74.0	7680	6	ABT11082	Abt11082 Human bre

ALIGNMENTS

RESULT 1
AAX81299
ID AAX81299 standard; cDNA; 2147 BP.
XX
AC AAX81299;
XX
DT 21-SEP-1999 (first entry)
XX
DE Human migration stimulating factor (MSF) 1-alpha encoding cDNA.
XX
KW Migration stimulatory factor; MSF; cell migration; modulation; human;
XX wound healing; scarring; MSF1-alpha; ss.
XX Homo sapiens.
XX
FN WO9931233-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-GB003766.
XX
PR 16-DEC-1997; 97GB-00026539.
XX
PA (UYDU-) UNIV DUNDEE.
XX
PI Schor SL, Schor AM;
XX
DR WPI; 1999-430039/36.
DR P-PSDB; AAY28901.

XX Proteins with cell migration stimulatory activity used in treating wound
PT and preventing scarring.
XX
PS Example 1; Fig 1; 86pp; English.
XX
CC The invention provides a human migration stimulatory factor (MSF)
CC protein. Host cells containing a replicable vector comprising the MSF
CC encoding nucleic acid can be used for the recombinant production of the
CC protein. The polypeptide can be used for modulating cell migration,
CC healing a wound and for preventing scarring. The present sequence
CC represents the nucleotide sequence encoding a human MSF1-alpha protein
XX
SQ Sequence 2147 BP; 575 A; 512 C; 575 G; 485 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.34e-12 Length: 2147
Score: 129.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-581-651D-7 (1-23) x AAX81299 (1-2147)

QY 1 ProCysValLeuProPheThrTyrAsnAspArgThrAspSerThrSerAsnTyrGlu 20
ADSI17488
ID ADS17488 standard; DNA; 1929 BP.
XX
AC ADS17488;
XX
XX
DT 02-DEC-2004 (first entry)
XX
DE Nucleotide sequence of human fibronectin 1.
XX
KW cell state; time-lapse profile; protein-protein interaction;
KW drug screening; cancer; infectious disease; allergy; hypertension;
KW hyperlipaemia; diabetes; cardiac disease; cerebral infarction; dementia;
KW obesity; arteriosclerosis; infertility; mental disease; nervous disease;
KW cataract; progeria; hypersensitivity; ultraviolet radiation; human;
KW fibronectin 1; actin acting substance; transfection array; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1929
FT /tag= a
FT /product= "fibronectin 1"
XX
XX WO2004079007-A2.
XX
XX 16-SEP-2004.
XX
XX 03-MAR-2004; 2004WO-JP002694.
XX
XX 04-MAR-2003; 2003JP-00057870.
XX
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX
XX Miyake M, Yoshikawa T, Uchimura E, Miyake J;
XX WPI; 2004-662438/64.
XX P-PSDB; ADS17489.
XX
XX Presenting a state of a cell, useful for diagnosing and treating a

PT disease, e.g. cancer, infectious disease, allergy, diabetes, dementia,
PT obesity, infertility, or cataract, comprises obtaining a time-lapse
XX profile of the cell.
PS Disclosure; SEQ ID NO 1; 532pp; English.
XX
CC The specification describes a method and system for accurately presenting
CC a state of a cell. The method comprises obtaining a time-lapse profile of
CC the cell by time-lapse monitoring of a gene state associated with at
CC least one gene derived from the cell, and presenting the time-lapse
CC profile. The gene comprises a transcription control sequence, and the
CC gene state includes expression of the gene. The method and system are
CC useful presenting a state of a cell. The method can allow the elucidation
CC of key protein-protein interactions suitable for targeting by drug
CC screening protocols. The method is useful for diagnosing or treating a
CC disease, e.g. cancer, infectious disease due to viruses or bacteria,
CC allergy, hypertension, hyperlipaemia, diabetes, cardiac disease, cerebral
CC infarction, dementia, obesity, arteriosclerosis, infertility, mental and
CC nervous diseases, cataract, progeria, or hypersensitivity to ultraviolet
CC radiation. The present sequence encodes human fibronectin 1. Bovine
CC fibronectin was used as a candidate for an actin acting substance. The
CC actin acting substance was used with transfection reagents and amplified
CC plasmid DNA in assays using transfection arrays, in the course of the
CC invention.
XX
SQ Sequence 1929 BP; 507 A; 462 C; 544 G; 416 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.95e-11 Length: 1929
Score: 122.00 Matches: 22
Percent Similarity: 95.65% Conservative: 0
Best Local Similarity: 95.65% Mismatches: 1
Query Match: 94.57% Indels: 0
DB: 13 Gaps: 0

US-09-581-651D-7 (1-23) x ADS17488 (1-1929)

QY 1 ProCysValLeuProPheThrTyrAsnAspArgThrAspSerThrSerAsnTyrGlu 20
ADSI17488
ID ADS17488 standard; DNA; 1929 BP.
XX
AC ADS17488;
XX
XX
DT 02-DEC-2004 (first entry)
XX
DE Human fibronectin 1 DNA, an actin acting substance SeqID 1.
XX
XX human; gene; ds; transfection efficiency; actin acting substance;
XX extracellular matrix; fibronectin 1; gene introduction reagent.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..1929
XX /tag= a
XX /product= "Fibronectin protein"
XX
XX WO2004079332-A2.
XX
XX 16-SEP-2004.
XX
XX 03-MAR-2004; 2004WO-JP002696.
XX
XX 04-MAR-2003; 2003JP-00057869.
XX
XX Presenting a state of a cell, useful for diagnosing and treating a

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 615.799 Seconds
(without alignments)
1809.796 Million cell updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO_spool_h/US09581651/runat_07112005_092223_28778/app_query.fasta_1.1834
-DB=genEmbl -QFMT=fractap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651.QCGN 1.1 9936 @runat_07112005_092223_28778 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -YGAPOP=6 -DELEXT=7
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.hcg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sv:*
13: gb.un:*
14: gb.vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	2147	6	BD137021 Polypepti
2	129	100.0	2147	6	AX003229 Sequence
3	129	100.0	2147	9	AJ276395 Homo sapi
4	122	94.6	1929	6	CQ871810 Sequence

ALIGNMENTS

RESULT 1
BD137021
LOCUS
DEFINITION Polypeptides, polynucleotides and uses thereof.
ACCESSION
VERSION BD137021.1 GI:23231966
KEYWORDS JP 2002508179-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2147)
Schor, S.L. and Schor, A.M.
AUTHORS
TITLE Polypeptides, polynucleotides and uses thereof
JOURNAL Patents: JP 2002508179-A 1 19-MAR-2002;
UNIVERSITY OF DUNDEE
COMMENT OS Homo sapiens (human)
PN JP 2002508179-A/1
PD 19-MAR-2002
PF 15-DEC-1998 JP 2000539133
PR 16-DEC-1997 GB 9726539.1
PI SETH LAWRENCE SCHOR, ANA MARIA SCHOR
PC C12N15/09, A61K38/00, A61P17/02, C07K14/78, C07K16/18, C12N5/10, PC C12P21/08,
PC C12O1/68, G01N33/574, C12N15/00, A61K37/02, C12N5/00 CC
Polypeptides, polynucleotides and uses thereof. FH Key
Location/Qualifiers 1. .2147
FT source

CQ871828 Sequence
CQ875358 Sequence
AX260622 Sequence
AF312399 Homo sapi
BD142777 Collagen-
CQ427970 Sequence
BD142785 Collagen-
BD133458 Collagen-
BD133469 Hybrid po
E63263 Collagen-bi
U42404 Human fibro
BD142784 Collagen-
E63271 Collagen-bi
BD133465 Collagen-
BD133476 Hybrid po
BD133466 Collagen-
E63270 Collagen-bi
BD133477 Hybrid po
AU535086 Homo sapi
CQ731571 Sequence
BC051082 Mus muscu
BX538045 Homo sapi
CQ715726 Sequence
BX640399 Homo sapi
BV178397 sqm10026
AL832771 Homo sapi
AL832202 Homo sapi
AR274901 Sequence
AR380744 Sequence
AX277596 Sequence
AX335368 Sequence
X02761 Human mRNa
A14133 Fibronectin
AR034630 Sequence
E01162 cDNA encodi
I70110 Sequence 16
AR364992 Sequence
AB191261 Homo sapi
AR051657 Sequence
BX640608 Homo sapi
BX640802 Homo sapi



GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 6.45098 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122

Sequence: 1 QWERTYLGALVCTCYGGR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	122	100.0	296	2	Q8C6J7	Q8C6J7 mus musculus
2	122	100.0	749	2	Q7Z391	Q7Z391 homo sapien
3	122	100.0	1103	2	Q6MZP4	Q6MZP4 homo sapien
4	122	100.0	2267	2	Q68DP9	Q68DP9 homo sapien
5	122	100.0	2295	2	Q6N0A6	Q6N0A6 homo sapien
6	122	100.0	2357	2	Q68DT4	Q68DT4 homo sapien
7	122	100.0	2386	1	FINC HUMAN	P02751 homo sapien
8	122	100.0	2444	2	Q6N0Z5	P11276 mus musculus
9	122	100.0	2477	1	FINC MOUSE	P04937 rattus norv
10	122	100.0	2477	2	Q6MZU5	Q6MZU5 homo sapien
11	122	100.0	2477	2	Q6MZU5	Q6MZU5 homo sapien
12	117	95.9	2265	1	FINC BOVIN	P07589 bos taurus
13	114	93.4	215	2	Q6DD34	Q6DD34 xenopus lae
14	114	93.4	2481	1	FINC XENLA	Q91740 xenopus lae
15	114	93.4	2481	2	Q6GQZ5	Q6GQZ5 xenopus lae
16	113	92.6	2240	2	Q68DP8	Q68DP8 homo sapien
17	84	68.9	922	2	Q93405	Q93405 brachydanio
18	84	68.9	2478	2	Q93406	Q93406 brachydanio
19	75	61.5	100	1	FINC RABIT	Q28749 oryctolagus
20	75	61.5	111	2	Q862C9	Q862C9 bos taurus
21	75	61.5	163	2	Q9H382	Q9H382 homo sapien
22	75	61.5	211	2	Q7L553	Q7L553 homo sapien
23	75	61.5	216	2	Q99KD0	Q99KD0 mus musculus
24	75	61.5	268	2	Q6FJES	Q6FJES homo sapien
25	75	61.5	358	2	Q7T152	Q7T152 brachydanio
26	75	61.5	412	2	Q71U44	Q71U44 bos taurus
27	75	61.5	810	2	Q6R3F3	Q6R3F3 mus musculus
28	75	61.5	1034	2	Q6N0B4	Q6N0B4 homo sapien
29	75	61.5	1255	2	Q6MZS0	Q6MZS0 homo sapien
30	75	61.5	1328	1	FINC PLEWA	Q91289 pleurodeles
31	75	61.5	2146	2	Q68CX6	Q68CX6 homo sapien

32 75 61.5 2193 2 Q6M2M7 Q6m2m7 homo sapien
33 67 54.9 2408 2 Q6JAN2 Q6jan2 brachydanio
34 58 47.5 351 2 O02816 O02816 oryctolagus
35 54 44.3 76 2 O46557 O46557 equus caball
36 49.5 40.6 332 2 Q9MC76 Q9mc76 peanut atun
37 49 40.2 156 2 Q70JQ4 Q70jq4 triticum ae
38 49 40.2 170 2 Q8HSL8 Q8hsl8 oryza sativ
39 48.5 39.8 197 2 Q7XJ70 Q7xmj0 oryza sativ
40 48 39.3 67 2 Q8MRE1 Q8mrel drosophila
41 48 39.3 554 1 NTP3 TOBAC P29162 nicotiana t
42 48 39.3 680 2 Q7PXK0 Q7pxm0 anopheles g
43 47.5 38.9 416 2 Q63L42 Q63l42 burkholderi
44 47 38.5 96 2 Q7YTR6 Q7ytr6 caenorhabdi
45 47 38.5 178 1 CRB2_BIFLO Q8g5c3 bifidobacte

ALIGNMENTS

RESULT 1
Q8C6J7 PRELIMINARY; PRT; 296 AA.
AC Q8C6J7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330027109 product:fibronectin 1, full insert sequence.
DE insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA The FANTOM Consortium,
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=20530913; PubMed=1076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kikuchi T., Taishiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

```
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Tanaka S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Iehii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tonaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK054456; BAC35784.1; -.
DR HSSP; Q96KP7; 1FBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR Pfam; PF00039; fnl; 5.
DR PRINTS; PR00012; FNTYPEI.
DR SMART; SM00058; FN1; 5.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 5.
SQ SEQUENCE 296 AA; 32617 MW; A0PF5F4809FB6439 CRC64;

Query Match 100.0%; Score 122; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSSR 21
Db 64 QWERTYLGNALVCTCYGSSR 84

RESULT 2
Q72391 ID Q72391 PRELIMINARY; PRT; 749 AA.
AC Q72391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -.
DR HSSP; Q96KP7; 1FBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fnl; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000395; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Tanaka S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Iehii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tonaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK054456; BAC35784.1; -.
DR HSSP; Q96KP7; 1FBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR Pfam; PF00039; fnl; 5.
DR PRINTS; PR00012; FNTYPEI.
DR SMART; SM00058; FN1; 5.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 5.
SQ SEQUENCE 296 AA; 32617 MW; A0PF5F4809FB6439 CRC64;

Query Match 100.0%; Score 122; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSSR 21
Db 64 QWERTYLGNALVCTCYGSSR 84

RESULT 2
Q72391 ID Q72391 PRELIMINARY; PRT; 749 AA.
AC Q72391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -.
DR HSSP; Q96KP7; 1FBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fnl; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000395; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.

DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 1103 AA; 82FEC4CAF634AD56 CRC64;

Query Match 100.0%; Score 122; DB 2; Length 1103;
Best Local Similarity 100.0%; Pred. No. 8.8e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSSR 21
Db 152 QWERTYLGNALVCTCYGSSR 172

RESULT 4
Q68DF9 ID Q68DF9 PRELIMINARY; PRT; 2267 AA.
AC Q68DF9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
```

```
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 100.0%; Score 122; DB 2; Length 749;
Best Local Similarity 100.0%; Pred. No. 6.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSSR 21
Db 155 QWERTYLGNALVCTCYGSSR 175

RESULT 3
Q6MZFA ID Q6MZFA PRELIMINARY; PRT; 1103 AA.
AC Q6MZFA;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219 (Fragment)).
GN Name=DKFZp686K139; Synonyms=DKFZp686F219;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RA The German Human cDNA Consortium;
RA Ansoerg W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CA546200.1; -.
DR EMBL; BX640802; CA545885.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR000957; FN_III-like.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 4.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 1103 AA; 82FEC4CAF634AD56 CRC64;

Query Match 100.0%; Score 122; DB 2; Length 1103;
Best Local Similarity 100.0%; Pred. No. 8.8e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSSR 21
Db 152 QWERTYLGNALVCTCYGSSR 172

RESULT 4
Q68DF9 ID Q68DF9 PRELIMINARY; PRT; 2267 AA.
AC Q68DF9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 1.59216 Seconds
(without alignments)
1269.066 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122

Sequence: 1 QWERTYLGNALVCTCGSSR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	100.0	2477	2	fibronectin precu
2	118	96.7	2386	1	fibronectin precu
3	117	95.9	2265	1	fibronectin - bovi
4	114	93.4	2481	2	fibronectin - Afri
5	75	61.5	103	2	fibronectin - mous
6	48	39.3	554	2	pollen-specific pr
7	47	38.5	210	2	protein B0416.3 [i
8	47	38.5	265	2	modulation protein
9	47	38.5	552	2	Y4NP protein - Rhi
10	47	38.5	774	2	protein tif9.8 [im
11	46	37.7	638	2	lichenan operon tr
12	46	37.7	655	1	hepatocyte growth
13	45.5	37.3	354	1	excisionase - Stap
14	45	36.9	106	2	GSA4 - Arabidopsi
15	45	36.9	106	2	gibberellin-regula
16	45	36.9	280	2	Glycoprotein - hum
17	45	36.9	300	1	nitrogen fixation
18	45	36.9	300	2	nitrogen fixation
19	45	36.9	335	2	cyanide insensitiv
20	45	36.9	436	2	adenosylmethionine
21	45	36.9	439	2	adenosylmethionine
22	44.5	36.5	98	2	30S ribosomal prot
23	44.5	36.5	409	2	sucrose cleavage p
24	44.5	36.5	1234	1	complement factor
25	44	36.1	65	2	probable membrane
26	44	36.1	149	2	plastoquinol-plast
27	44	36.1	190	2	fibronectin - east
28	44	36.1	276	2	maltose transport
29	44	36.1	297	2	probable hydrolase

30 44 36.1 818 2 F96586
31 43.5 35.7 303 2 H35068
32 43.5 35.7 452 2 A35068
33 43.5 35.7 808 2 D35069
34 43.5 35.7 822 2 AB0238
35 43.5 35.7 822 2 T47007
36 43.5 35.7 2810 2 T22298
37 43 35.2 135 2 B28823
38 43 35.2 160 2 JQ0542
39 43 35.2 171 2 S57894
40 43 35.2 189 2 B69044
41 43 35.2 220 2 T45056
42 43 35.2 250 2 A34342
43 43 35.2 345 2 H95267
44 43 35.2 426 2 T04985
45 43 35.2 507 2 T38653

ALIGNMENTS

RESULT 1

S14428

fibronectin precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text change 09-Jul-2004

C:Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049

R:Hynes, R.O.

submitted to the EMBL Data Library, July 1989

A:Reference number: S14428

A:Accession: S14428

A:Molecule type: mRNA

A:Residues: 1-2477 <HYN>

A:Cross-references: UNIPROT:P04937; EMBL:X15906; NID:G56163; PIDN:CAA34020.1; PID:G56164

R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.

EMBO J. 6, 2573-2580, 1987

A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.

A:Reference number: S12455; MUID:88054951; PMID:2445560

A:Accession: S12455

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 609-1810, 'T', 1812-2283 <SCH>

A:Cross-references: EMBL:X15906

R:Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.

Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984

A:Title: A single rat fibronectin gene generates three different mRNAs by alternative sp

A:Reference number: A22319; MUID:84298097; PMID:6089177

A:Accession: A22319

A:Molecule type: DNA

A:Residues: 2052-2237 <TAM>

R:Palkenberg, C.; Enghild, J.J.; Thogersen, I.B.; Salvesen, G.; Akerstroem, B.

Biochem. J. 301, 745-751, 1994

A:Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex in

A:Reference number: S46203; MUID:94330948; PMID:7519849

A:Accession: S46203

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1183-1192; 'GUN', 1268, 'P', 1270-1271, 'D', 1273, 'CF', 1276, 'PV', 1385-1399 <PAL>

R:Patel, R.S.; Odermatt, E.; Schwarzbauer, J.E.; Hynes, R.O.

EMBO J. 6, 2565-2572, 1987

A:Title: Organization of the fibronectin gene provides evidence for exon shuffling durin

A:Reference number: S00459; MUID:88054950; PMID:3119323

A:Accession: S00459

A:Molecule type: DNA

A:Residues: 1-139; 2382-2477 <PAT>

A:Cross-references: EMBL:X05831

A:Note: the authors translated the codon CCT for residues 51 and 94 as Ala

R:Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hynes, R.O.

Cell 35, 421-431, 1983

A:Title: Three different fibronectin mRNAs arise by alternative splicing within the codi

A:Reference number: A27252; MUID:84082067; PMID:6317187

A:Accession: A27252

A:Molecule type: mRNA

A;Residues: 1586-1720,'T',1722,1813-2477 <SC2>
 R;Odermatt, E.; Tankun, J.W.; Hynes, R.O.
 Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985
 A;Title: Repeating modular structure of the fibronectin gene: Relationship to protein structure
 A;Reference number: 159049; MUID:86016741; PMID:3863113
 A;Accession: 159049
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1722-1810 <RES>
 A;Cross-references: GB:M11750; NID:g204164; PIDN:AAA41170.1; PID:g554437
 C;Genetics:
 A;Introns: 51/1; 94/1; 2416/3; 2454/3
 C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat homology
 C;Keywords: alternative splicing; cell adhesion; collagen binding; disulfide bond; duplication
 F;1-32/Domain: signal sequence #status predicted <SIG>
 F;33-2477/Product: fibronectin #status predicted <MAT>
 F;53-88/Domain: fibronectin type I repeat homology <1F1>
 F;98-136/Domain: fibronectin type I repeat homology <1F2>
 F;142-180/Domain: fibronectin type I repeat homology <1F3>
 F;187-226/Domain: fibronectin type I repeat homology <1F4>
 F;232-271/Domain: fibronectin type I repeat homology <1F5>
 F;308-342/Domain: fibronectin type I repeat homology <1F6>
 F;360-401/Domain: fibronectin type II repeat homology <2F1>
 F;420-461/Domain: fibronectin type II repeat homology <2F2>
 F;470-508/Domain: fibronectin type I repeat homology <1F7>
 F;518-555/Domain: fibronectin type I repeat homology <1F8>
 F;561-599/Domain: fibronectin type I repeat homology <1F9>
 F;609-692/Domain: fibronectin type III repeat homology <FN3A>
 F;718-800/Domain: fibronectin type III repeat homology <FN3B>
 F;809-890/Domain: fibronectin type III repeat homology <FN3C>
 F;905-987/Domain: fibronectin type III repeat homology <FN3D>
 F;995-1076/Domain: fibronectin type III repeat homology <FN3E>
 F;1085-1164/Domain: fibronectin type III repeat homology <FN3F>
 F;1172-1257/Domain: fibronectin type III repeat homology <FN3G>
 F;1265-1348/Domain: fibronectin type III repeat homology <FN3H>
 F;1356-1439/Domain: fibronectin type III repeat homology <FN3I>
 F;1447-1529/Domain: fibronectin type III repeat homology <FN3J>
 F;1537-1619/Domain: fibronectin type III repeat homology <FN3K>
 F;1634-1616/Region: cell attachment (R-G-D) motif
 F;1631-1713/Domain: fibronectin type III repeat homology <FN3L>
 F;1721-1803/Domain: fibronectin type III repeat homology <FN3M>
 F;1811-1893/Domain: fibronectin type III repeat homology <FN3N>
 F;1903-1984/Domain: fibronectin type III repeat homology <FN3O>
 F;1992-2074/Domain: fibronectin type III repeat homology <FN3P>
 F;2181-2183/Region: cell attachment (R-G-D) motif
 F;2193-2273/Domain: fibronectin type III repeat homology <FN3Q>
 F;2296-2335/Domain: fibronectin type I repeat homology <1F10>
 F;2341-2378/Domain: fibronectin type I repeat homology <1F11>
 F;2385-2420/Domain: fibronectin type I repeat homology <1F12>
 F;53-79,77-88,98-126,124-136,142-170,168-180,187-216,214-226,232-261,259-271,308-335,333-368,2366-2378,2385-2411,2409-2420/Disulfide bonds: #status predicted
 F;2458/Disulfide bonds: interchain (to 2462) #status predicted
 F;2462/Disulfide bonds: interchain (to 2458) #status predicted

Query Match 100.0%; Score 122; DB 2; Length 2477;
 Best Local Similarity 100.0%; Pred.No. 7.4e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQWERTYLGNALVCTCYGGSR 21
 Db 64 QQWERTYLGNALVCTCYGGSR 84

RESULT 2
 FNHU
 N;Alternate names: fibronectin splice form ED-A
 C;Species: Homo sapiens (man)
 C;Date: 27-Nov-1985 #sequence_revision 31-Mar-1993 #text change 09-Jul-2004
 A;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22
 R;Dean, D.C.; Bowlus, C.L.; Bourgeois, S.
 Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
 A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A;Reference number: A26460; MUID:87175578; PMID:3031656
 A;Accession: A26460
 A;Molecule type: DNA
 A;Residues: 1-49 <DEA>
 A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAA5337
 R;Oldberg, A.; Ruoslahti, E.
 J. Biol. Chem. 261, 2113-2116, 1986
 A;Title: Evolution of the fibronectin gene.
 A;Reference number: A26284; MUID:86111901; PMID:3003095
 A;Accession: A26284
 A;Molecule type: DNA
 A;Residues: 1447-1540 <OLD>
 A;Cross-references: GB:M12549; NID:g182688
 A;Note: the authors translated the codon TTC for residue 1494 as Glu
 R;Paoletta, G.; Henchcliffe, C.; Sebastiao, G.; Baralle, F.E.
 Nucleic Acids Res. 16, 3545-3557, 1988
 A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B
 A;Reference number: S00848; MUID:88233940; PMID:3375063
 A;Accession: S03917
 A;Molecule type: DNA
 A;Residues: 1594-1767,'V',1769-1783 <PAO>
 A;Cross-references: EMBL:X07718; NID:g31402
 A;Note: the authors translated the codon AAC for residue 1631 as Asp
 R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
 PNAS Lett. 207, 287-291, 1986
 A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene:
 A;Reference number: A24854; MUID:87030929; PMID:3770201
 A;Accession: A24854
 A;Molecule type: DNA
 A;Residues: 1992-2147 <VIB>
 A;Cross-references: GB:X04530; NID:g31436
 R;Gutman, A.; Yamada, K.M.; Kornblihtt, A.
 FEBS Lett. 207, 145-148, 1986
 A;Title: Human fibronectin is synthesized as a pre-propolypeptide.
 A;Reference number: A24476; MUID:87030890; PMID:3770189
 A;Accession: A24476
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-14,'Q',16-38 <GUT>
 R;Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
 EMBO J. 4, 1755-1759, 1985
 A;Title: Primary structure of human fibronectin: differential splicing may generate at l
 A;Reference number: A91008; MUID:85284965; PMID:2992939
 A;Accession: A91008
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 32-1344,1346-2080;2112-2386 <KOR>
 A;Cross-references: GB:X02761
 R;Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Nucleic Acids Res. 12, 5853-5868, 1984
 A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
 A;Reference number: A93529; MUID:84272258; PMID:6462919
 A;Accession: A93529
 A;Molecule type: mRNA
 A;Residues: 973-2080;2112-2386 <KO2>
 A;Cross-references: GB:X00739
 R;Oldberg, A.; Linney, E.; Ruoslahti, E.
 J. Biol. Chem. 258, 10193-10196, 1983
 A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell
 A;Reference number: A21011; MUID:83290929; PMID:6688418
 A;Accession: A21011
 A;Molecule type: mRNA
 A;Residues: 1434-1537 <OL2>
 A;Cross-references: GB:X00055; NID:g182680; PIDN:AAA52459.1; PID:g182683
 R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
 Biochemistry 24, 2698-2704, 1985
 A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with r
 A;Reference number: A90495; MUID:85280409; PMID:2992573
 A;Accession: A90495
 A;Molecule type: mRNA
 A;Residues: 1594-2386 <BER>
 A;Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697
 R;Umezawa, K.; Kornblihtt, A.R.; Baralle, F.E.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 7.1098 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122

Sequence: 1 QQWERTYLGALVCTCYGGR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122	100.0	231	9	US-09-925-302-548
2	122	100.0	231	10	Sequence 548, App
3	122	100.0	463	15	US-09-925-302-548
4	122	100.0	463	16	Sequence 52, App
5	122	100.0	642	16	US-10-144-194A-52
6	122	100.0	642	17	US-10-491-566-52
7	122	100.0	657	16	US-10-741-601-354
8	122	100.0	657	17	US-10-741-601-1066
9	122	100.0	657	17	US-10-741-601-359
10	122	100.0	984	16	US-10-741-600-1072
11	122	100.0	984	17	US-10-741-601-356
12	122	100.0	1173	18	US-10-741-600-1069
13	122	100.0	1173	18	US-10-450-763-52634

12	122	100.0	2220	15	US-10-236-392-4	Sequence 4, Appl
13	122	100.0	2296	16	US-10-741-601-363	Sequence 363, App
14	122	100.0	2296	17	US-10-741-600-1075	Sequence 1075, App
15	122	100.0	2320	14	US-10-279-733-8	Sequence 8, Appl
16	122	100.0	2325	15	US-10-144-194A-104	Sequence 104, App
17	122	100.0	2355	15	US-10-360-101-235	Sequence 235, App
18	122	100.0	2355	15	US-10-447-161-3	Sequence 3, Appl
19	122	100.0	2355	16	US-10-734-564-94	Sequence 94, Appl
20	122	100.0	2355	16	US-10-741-601-366	Sequence 366, App
21	122	100.0	2355	16	US-10-741-601-366	Sequence 367, App
22	122	100.0	2355	16	US-10-491-566-104	Sequence 104, App
23	122	100.0	2355	17	US-10-741-600-1067	Sequence 1067, App
24	122	100.0	2355	17	US-10-741-600-1078	Sequence 1078, App
25	122	100.0	2355	17	US-10-852-335A-147	Sequence 147, App
26	122	100.0	2355	18	US-10-287-436A-436	Sequence 436, App
27	122	100.0	2355	18	US-10-287-436A-1137	Sequence 1137, App
28	122	100.0	2355	20	US-11-040-130-28	Sequence 28, Appl
29	122	100.0	2386	10	US-09-961-403-1	Sequence 1, Appl
30	122	100.0	2386	16	US-10-741-601-360	Sequence 360, App
31	122	100.0	2386	16	US-10-788-792-206	Sequence 206, App
32	122	100.0	2386	16	US-10-618-281-32	Sequence 32, Appl
33	122	100.0	2386	17	US-10-741-600-1071	Sequence 1071, App
34	122	100.0	2386	17	US-10-868-577A-59	Sequence 59, Appl
35	122	100.0	2386	18	US-10-485-758-4	Sequence 4, Appl
36	122	100.0	2386	18	US-10-485-758-9	Sequence 9, Appl
37	118	96.7	259	10	US-09-940-235-4	Sequence 4, Appl
38	118	96.7	2320	15	US-10-236-392-2	Sequence 2, Appl
39	118	96.7	2328	14	US-10-171-311-64	Sequence 64, Appl
40	118	96.7	2328	15	US-10-236-031B-70	Sequence 70, Appl
41	118	96.7	2328	15	US-10-374-979-98	Sequence 98, Appl
42	118	96.7	2328	15	US-10-182-936A-98	Sequence 98, Appl
43	118	96.7	2328	16	US-10-477-238A-677	Sequence 677, App
44	118	96.7	2328	16	US-10-680-287A-677	Sequence 677, App
45	118	96.7	2328	17	US-10-477-173-677	Sequence 677, App

ALIGNMENTS

RESULT 1

US-09-925-302-548
; Sequence 548, Application US/09925302
; Patent No. US2002004941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P4104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 548
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (212)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (226)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-302-548

```

RESULT 6
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: MICHELE et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741.600

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 2.22353 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-6
Perfect score: 122
Sequence: 1 QWERTYLGALVCTCYGSR 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	100.0	2386	4	US-09-961-403-1
2	118	96.7	2231	1	US-08-153-799-16
3	118	96.7	2324	1	US-08-283-857-1
4	118	96.7	2324	5	PCT-US95-09819-1
5	118	96.7	2327	6	5455158-1
6	118	96.7	2327	6	5455158-1
7	118	96.7	2386	2	US-09-016-366A-12
8	118	96.7	2446	2	US-08-551-356-2
9	118	96.7	2446	5	PCT-US93-12687-2
10	75	61.5	188	1	US-08-142-449B-14
11	75	61.5	422	2	US-08-836-854-12
12	75	61.5	474	2	US-08-836-854-9
13	54	44.3	42	1	US-08-340-812-7
14	54	44.3	42	1	US-08-459-064B-7
15	54	44.3	42	1	US-08-460-421A-7
16	54	44.3	42	5	PCT-US93-00909-7
17	50.5	41.4	48	2	US-08-982-597A-20
18	50.5	41.4	48	3	US-09-136-218-20
19	47	38.5	668	4	US-09-248-796A-18675
20	46	37.7	276	4	US-09-270-767-43364
21	46	37.7	655	1	US-08-148-910-12
22	46	37.7	655	1	US-08-448-937A-12
23	45	36.9	37	2	US-08-982-597A-17
24	45	36.9	37	3	US-09-136-218-17
25	45	36.9	336	4	US-09-252-991A-20404
26	44	36.1	123	3	US-08-793-450-4
27	44	36.1	353	4	US-09-543-681A-5307

Sequence 14, Appl
Sequence 13133, A
Sequence 18, Appl
Sequence 18, Appl
Sequence 24, Appl
Sequence 7031, Ap
Sequence 2, Appl
Sequence 5150, Ap
Sequence 21, Appl
Sequence 5158, Ap
Sequence 1560, Ap
Sequence 7975, Ap
Sequence 12, Appl
Sequence 12, Appl
Sequence 12, Appl

28 43 35.2 227 1 US-07-869-933-14
29 43 35.2 227 3 US-09-103-663-14
30 43 35.2 544 4 US-09-489-039A-13133
31 43 35.2 587 2 US-08-871-266B-18
32 43 35.2 587 3 US-09-018-864A-18
33 43 35.2 587 3 US-08-871-267B-24
34 43 35.2 587 3 US-09-618-419-24
35 43 35.2 593 4 US-09-949-016-7031
36 43 35.2 2321 4 US-09-230-652-2
37 42.5 34.8 232 4 US-09-328-352-5150
38 42 34.4 43 2 US-08-982-597A-21
39 42 34.4 43 3 US-09-136-218-21
40 42 34.4 56 4 US-09-513-999C-5158
41 42 34.4 143 4 US-09-732-210-1560
42 42 34.4 143 4 US-09-513-999C-7975
43 42 34.4 164 4 US-09-907-794A-12
44 42 34.4 164 4 US-09-905-125A-12
45 42 34.4 164 4 US-09-902-775A-12

ALIGNMENTS

RESULT 1
US-09-961-403-1
; Sequence 1, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961.403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-1

Query Match 100.0%; Score 122; DB 4; Length 2386;
Best Local Similarity 100.0%; Pred. No. 9.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGALVCTCYGSR 21
Db 63 QWERTYLGALVCTCYGSR 83

RESULT 2
US-08-153-799-16
; Sequence 16, Application US/08153799
; Patent No. 5768883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESS: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2231 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Protein
LOCATION: 1..2231
OTHER INFORMATION: /notes="Human fibronectin"
US-08-153-799-16

Query Match 96.7%; Score 118; DB 1; Length 2231;
Best Local Similarity 95.2%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGGSR 21
Db 32 QWERTYLGNVLVCTCYGGSR 52

RESULT 3
US-08-283-857-1
Sequence 1, Application US/08283857
Patent No. 5792742
GENERAL INFORMATION:
APPLICANT: GOLD, Leslie I.
APPLICANT: ROSTAGNO, Agueda A.
APPLICANT: BARON, Martin
APPLICANT: CAMPBELL, Iain D.
APPLICANT: WILLIAMS, Michael, J.
TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/283,857
FILING DATE: 01-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/714,134
FILING DATE: 14-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: GOLD-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2324 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-283-857-1

Query Match 96.7%; Score 118; DB 1; Length 2324;
Best Local Similarity 95.2%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGGSR 21
Db 32 QWERTYLGNVLVCTCYGGSR 52

RESULT 4
PCT-US95-09819-1
Sequence 1, Application PC/TUS9509819
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09819
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,857
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: GOLD-1A PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2324 amino acids

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 7.76863 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122

Sequence: 1 QWERTYLGALVCTCYGGR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq_16Dec04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	100.0	21	2	AAY28903 MSF 1-alp
2	122	100.0	231	3	AAB58210 Lung canc
3	122	100.0	463	6	ABR58303 BC00770 p
4	122	100.0	642	8	ADQ39403 Human myo
5	122	100.0	642	8	ADR67316 Human bla
6	122	100.0	642	8	ADS17489 Amino aci
7	122	100.0	642	8	ADR97658 Human fib
8	122	100.0	657	8	ADQ39409 Human myo
9	122	100.0	660	2	AAY28901 Human mig
10	122	100.0	984	8	ADQ39406 Human myo
11	122	100.0	1173	4	ABG22275 Novel hum
12	122	100.0	1223	8	ADP30365 Human sec
13	122	100.0	2182	8	ADR66462 Human pro
14	122	100.0	2182	8	ADR66120 Human pro
15	122	100.0	2220	6	ABO01289 Human pro
16	122	100.0	2220	8	ADN95950 Human NOV
17	122	100.0	2225	4	RAM38647 Human pol
18	122	100.0	2266	6	ABR40124 Human cel
19	122	100.0	2296	8	ADQ39412 Human myo
20	122	100.0	2320	6	RAE37107 Human fib
21	122	100.0	2320	6	ABR42588 Human fib
22	122	100.0	2330	4	RAM38646 Human pol
23	122	100.0	2355	4	RAM38649 Human pol
24	122	100.0	2355	6	ABR58335 NM_00202
25	122	100.0	2355	7	ADP65196 Human fib

26	122	100.0	2355	8	ADG89560 Human fib
27	122	100.0	2355	8	ADL92160 Fibronect
28	122	100.0	2355	8	ADO55175 Protein #
29	122	100.0	2355	8	ADQ26085 Fibronect
30	122	100.0	2355	8	ADQ29668 Human col
31	122	100.0	2355	8	ADQ39415 Human myo
32	122	100.0	2355	8	ADQ39404 Human myo
33	122	100.0	2355	8	ADR67315 Human bla
34	122	100.0	2386	4	AAM38648 Human pol
35	122	100.0	2386	5	AAO17353 Human fib
36	122	100.0	2386	6	ABR81866 Human fib
37	122	100.0	2386	7	ADD18770 Human dis
38	122	100.0	2386	7	ADE63324 Human Pro
39	122	100.0	2386	8	ADO55174 Protein #
40	122	100.0	2386	8	ADQ39408 Human myo
41	122	100.0	2386	8	ADR99200 Fibronect
42	122	100.0	2446	3	AAB50377 Human fib
43	122	100.0	2447	4	AAM40434 Human pol
44	122	100.0	2447	4	AAM40435 Human pol
45	122	100.0	2447	4	AAM40433 Human pol

ALIGNMENTS

RESULT 1

AAY28903

ID AAY28903 standard; peptide; 21 AA.

AC AAY28903;

DT 21-SEP-1999 (first entry)

DE MSF 1-alpha peptide epitope.

XX Migration stimulatory factor; MSF; cell migration; modulation; human;
wound healing; scarring; MSF1-alpha; epitope; fibronectin.

OS Synthetic.

OS Homo sapiens.

PN WO9931233-A1.

PD 24-JUN-1999.

PF 15-DEC-1998; 98WO-GB003766.

PR 16-DEC-1997; 97GB-00026539.

PA (UYDU-) UNIV DUNDEE.

PI Schor SL, Schor AM;

DR WPI; 1999-430039/36.

PT Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX Claim 18; Page 57; 86pp; English.

CC The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. Sequences AAY28902-906 represent peptide epitopes of MSF against which monoclonal antibodies that are specific to MSF and do not cross-react with fibronectin are raised

XX Sequence 21 AA;

SQ

Query Match 100.0%; Score 122; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.8e-09;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSR 21
 Db 1 QWERTYLGNALVCTCYGSR 21

RESULT 2
 AAB58210
 ID AAB58210 standard; protein; 231 AA.
 AC AAB58210;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 548.
 XX
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active; vulnerary;
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200055180-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005918.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Ruben SM;
 XX
 WPI; 2000-587514/55.
 DR N-PSDB; AAF18086.
 XX
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer.
 XX
 PS Claim 11; Page 1041-1042; 1425pp; English.
 XX
 CC Polynucleotide sequences AAF1982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytostatic; cardioactive;
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the protein
 CC or polynucleotide sequences. The lung cancer associated polynucleotide
 CC sequences may be used for detection of lung cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The proteins may be used to treat disorders such as
 CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
 CC cardiovascular, renal, and proliferative disorders. The proteins may also
 CC be used in the treatment of wounds and infectious diseases.
 CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
 CC used in the course of the invention for the identification and
 CC characterization of the polynucleotide and protein sequences
 XX
 SQ Sequence 231 AA;

Query Match 100.0%; Score 122; DB 3; Length 231;
 Best Local Similarity 100.0%; Pred. No. 1.8e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSR 21
 Db 1 QWERTYLGNALVCTCYGSR 21

RESULT 4
 ADQ39403
 ID ADQ39403 standard; protein; 642 AA.
 XX
 AC ADQ39403;
 XX
 DT 18-NOV-2004 (first entry)
 XX

Db 92 QWERTYLGNALVCTCYGSR 112

RESULT 3
 ABR58303
 ID ABR58303 standard; protein; 463 AA.
 XX
 AC ABR58303;
 XX
 DT 08-JUL-2003 (first entry)
 XX
 DB BCU0770 protein #SEQ ID 52.
 XX
 KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
 KW drug discovery; clinical medicine; forensic medicine; chromosome 2q35.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029421-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-US031287.
 XX
 PR 03-OCT-2001; 2001US-0326526P.
 PR 14-MAY-2002; 2002US-00144194.
 XX
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX
 PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
 XX
 WPI; 2003-381623/36.
 DR N-PSDB; ACC72037.
 XX
 PT New isolated human differentially-regulated breast cancer polynucleotide
 PT and polypeptide, useful for diagnosing, staging, prognosticating,
 PT preventing and/or treating diseases and conditions relating to breast
 PT cancer.
 XX
 PS Claim 6; SEQ ID NO 52; 127pp + Sequence Listing; English.
 XX
 CC The invention relates to isolated polynucleotides which are
 CC differentially-regulated in breast cancer. The methods and compositions
 CC of the present invention are useful for detecting, diagnosing, staging,
 CC monitoring, prognosticating, preventing and/or treating diseases and
 CC conditions relating to breast cancer, and may be used in gene therapy or
 CC antisense therapy. They can also be used in research, drug discovery,
 CC clinical medicine and forensic medicine. Sequences given in records
 CC ABR58278-ABR58323 and ABR58346-ABR58362 represent polypeptides encoded by
 CC polynucleotides of the invention that are differentially-regulated in
 CC breast cancer. NOTE: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 463 AA;

Query Match 100.0%; Score 122; DB 6; Length 463;
 Best Local Similarity 100.0%; Pred. No. 3.5e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSR 21
 Db 99 QWERTYLGNALVCTCYGSR 119

RESULT 4
 ADQ39403
 ID ADQ39403 standard; protein; 642 AA.
 XX
 AC ADQ39403;
 XX
 DT 18-NOV-2004 (first entry)
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 436.745 Seconds
(without alignments)
1830.243 Million cell updates/sec

Title: US-09-581-651D-6
Perfect score: 122
Sequence: 1 QQWERTYLGALVCTCYGSSR 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPRO_spool_b/US09581651/runat_07112005_092223_28789/app_query.fasta_1.1834
-DB=EST -QFMT=fascap -SUPFIX=p2n.rst -MINMATCH=0.1 -LOOPC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651_QCGN 1.1 8180 @runat_07112005_092223_28789 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_hcc.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gsl1.*

9: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

c 1	122	100.0	239	2	BF358738 CM2-ET001
2	122	100.0	292	4	BF988947 IL5-GN017
3	122	100.0	305	2	AW068290 cn20c07.Y
4	122	100.0	317	9	CU212775 G050E08 G
5	122	100.0	321	2	AW068245 cn23g07.Y
6	122	100.0	343	2	B8817493 QV4-BN026
7	122	100.0	349	6	CB704303 AMGNNUC:C
c 8	122	100.0	378	2	BF751799 RC3-BN042
9	122	100.0	407	9	CG536337 OST124189

c 10	122	100.0	427	4	BF988908
11	122	100.0	435	1	AU140826
c 12	122	100.0	446	2	BF329480
13	122	100.0	465	7	CF166902
14	122	100.0	473	7	CF166406
15	122	100.0	480	4	BF993488
16	122	100.0	481	7	CF896895
c 17	122	100.0	487	4	BF993485
c 18	122	100.0	491	4	BF993489
19	122	100.0	494	5	BQ366579
20	122	100.0	494	5	BQ366579
21	122	100.0	514	1	AL710256
22	122	100.0	516	5	BA509989
23	122	100.0	521	5	BA509989
24	122	100.0	525	5	BA509989
25	122	100.0	528	5	BA509989
26	122	100.0	530	1	AL702029
27	122	100.0	530	5	BA509989
28	122	100.0	532	5	BA509989
29	122	100.0	535	1	AL699665
30	122	100.0	536	5	BA509989
31	122	100.0	537	5	BA509989
32	122	100.0	539	5	BA509989
33	122	100.0	540	5	BA509989
34	122	100.0	541	5	BA509989
35	122	100.0	546	5	BA509989
36	122	100.0	548	5	BA509989
37	122	100.0	548	5	BA509989
38	122	100.0	548	7	CN332331
39	122	100.0	554	5	BA509989
40	122	100.0	561	5	BA509989
41	122	100.0	562	5	BA509989
42	122	100.0	565	1	AL702074
43	122	100.0	565	1	AL702089
44	122	100.0	565	1	AL712567
45	122	100.0	565	5	BA508450

ALIGNMENTS

RESULT 1
BF358738/c

LOCUS BF358738 239 bp mRNA linear EST 22-NOV-2000

DEFINITION CM2-ET0016-310500-206-f05 ET0016 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF358738

VERSION BF358738.1 GI:11317810

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 239)

AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Negati, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brundstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,

O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-ET0016-310500-206-f05&t3=2000-05-31&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 10
 High quality sequence stop: 239.
 Location/Qualifiers

FEATURES

source
 1..239
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="ET0016"
 /note="Organ: lung tumor; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 1.13e-08 Length: 239
 Score: 122.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-581-651D-6 (1-21) x BF358738 (1-239)

Qy 1 GlnGlnTrpGluArgThrTyLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
 |||||
 Db 190 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGGAGGAAGC 131

Qy 21 Arg 21
 |||
 Db 130 CGA 128

RESULT 2

BF988947
 LOCUS IL5-GN0179-181000-176-c08 GN0179 Homo sapiens cDNA, mRNA sequence. EST 23-JAN-2001

ACCESSION BF988947

VERSION BF988947.1 GI:12395272

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-GN0179-310500-206-f05&t3=2000-05-31&t4=1>)

181000-176-c08&t3=2000-10-18&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 291.
 Location/Qualifiers

FEATURES

source
 1..292
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="GN0179"
 /note="Organ: placenta normal; Vector: puc18; Site 1:
 SmaI; Site 2: SmaI; A mini-library was made by cloning
 products derived from ORESTES PCR (U.S. Letters Patent
 Application No. 196,716 - Ludwig Institute for Cancer
 Research) profiles into the puc 18 vector. Reverse
 transcription of tissue mRNA and cDNA amplification were
 performed under low stringency conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 1.43e-08 Length: 292
 Score: 122.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-581-651D-6 (1-21) x BF988947 (1-292)

Qy 1 GlnGlnTrpGluArgThrTyLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
 |||||
 Db 193 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGGAGGAAGC 252

Qy 21 Arg 21
 |||
 Db 253 CGA 255

RESULT 3

AW068290
 LOCUS cn20c07.y1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
 DEFINITION clone NHTBC_cn20c07 random, mRNA sequence.

ACCESSION AW068290

VERSION AW068290.1 GI:6023288

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G., Yang,L.M.,
 Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
 Robey,P.G., Hotchkiss,R.N. and Franccomano,C.A.

TITLE SGAP: The Skeletal Genome Anatomy Project

JOURNAL Unpublished (1997)

COMMENT Contact: Libin Jia

Medical Genetics Branch

National Human Genome Research Institute

10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA

Tel: 301-402-4877

Fax: 301-496-7157

Email: libin@helix.nih.gov

DNA Sequencing and analyses by National Institutes of Health
 Intramural Sequencing Center (NISC).

Plate: 20 row: c column: 07

Seq primer: M13RP1 reverse primer (ABI).

FEATURES

source
 1..305
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NHTBC_cn20c07"
 /sex="Female"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 107.827 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-6
Perfect score: 122
Sequence: 1 QQWERTYLGALVCTCYGGR 21

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1834
-DB=Published Applications NA -Qfmt=fastap -Suffix=p2n.rnpb -MINMATCH=0.1
-TRANSPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -USER=US09581651 @CGN 1.1 1041 @runat_07112005_092226_28867
-NCPU=6 -ICPU=3 -NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRSADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	122	100.0	407	9	US-09-728-445-799	Sequence 799, App
2	122	100.0	407	24	US-10-964-549-799	Sequence 799, App
3	122	100.0	867	9	US-09-925-302-105	Sequence 105, App
4	122	100.0	867	10	US-09-925-302-105	Sequence 105, App
5	122	100.0	2127	17	US-10-210-120-49	Sequence 49, Appl
6	122	100.0	2127	22	US-10-956-157-4288	Sequence 4288, Ap
7	122	100.0	2127	24	US-10-909-035-49	Sequence 49, Appl
8	122	100.0	2443	20	US-10-741-601-70	Sequence 70, Appl
9	122	100.0	2443	22	US-10-741-600-238	Sequence 238, App
10	122	100.0	2488	20	US-10-741-601-75	Sequence 75, Appl
11	122	100.0	3522	24	US-10-450-763-22266	Sequence 244, App
12	122	100.0	3522	24	US-10-450-763-22266	Sequence 244, App
13	122	100.0	4295	18	US-10-144-194A-51	Sequence 51, Appl
14	122	100.0	4295	21	US-10-491-566-51	Sequence 51, Appl
15	122	100.0	6510	20	US-10-741-601-72	Sequence 72, Appl
16	122	100.0	6510	22	US-10-741-600-241	Sequence 241, App
17	122	100.0	7361	19	US-10-236-392-3	Sequence 3, Appli
18	122	100.0	7677	22	US-10-956-157-4995	Sequence 4995, Ap
19	122	100.0	7795	16	US-10-084-817-2	Sequence 2, Appli
20	122	100.0	7823	20	US-10-741-601-77	Sequence 77, Appl
21	122	100.0	7823	22	US-10-741-600-245	Sequence 245, App
22	122	100.0	7848	20	US-10-741-601-78	Sequence 78, Appl
23	122	100.0	7848	22	US-10-741-600-246	Sequence 246, App
24	122	100.0	7867	14	US-10-098-841-6	Sequence 6, Appli
25	122	100.0	7935	20	US-10-741-601-74	Sequence 74, Appl
26	122	100.0	7935	22	US-10-741-600-240	Sequence 240, App
27	122	100.0	7959	20	US-10-741-601-81	Sequence 81, Appl
28	122	100.0	7959	22	US-10-741-600-249	Sequence 249, App
29	122	100.0	8013	20	US-10-741-601-71	Sequence 71, Appl
30	122	100.0	8013	22	US-10-741-600-242	Sequence 242, App
31	122	100.0	8027	18	US-10-447-161-8	Sequence 8, Appli
32	122	100.0	8027	20	US-10-734-564-27	Sequence 27, Appl
33	122	100.0	8027	22	US-10-852-335A-53	Sequence 53, Appl
34	122	100.0	8027	24	US-10-287-436A-81	Sequence 81, Appl
35	122	100.0	8044	17	US-10-240-965-121	Sequence 121, App
36	122	100.0	8044	24	US-10-765-700-135	Sequence 135, App
37	122	100.0	8062	14	US-10-098-841-5	Sequence 5, Appli
38	122	100.0	8137	14	US-10-098-841-8	Sequence 8, Appli
39	122	100.0	8155	20	US-10-741-601-79	Sequence 79, Appl
40	122	100.0	8155	22	US-10-741-600-247	Sequence 247, App
41	122	100.0	8226	20	US-10-741-601-69	Sequence 69, Appl
42	122	100.0	8226	22	US-10-741-600-237	Sequence 237, App
43	122	100.0	8230	14	US-10-098-841-7	Sequence 7, Appli
44	122	100.0	8278	20	US-10-741-601-82	Sequence 82, Appl
45	122	100.0	8278	22	US-10-741-600-250	Sequence 250, App

ALIGNMENTS

RESULT 1
US-09-728-445-799
; Sequence 799, Application US/09728445
; Patent No. US20020102543A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020102543A1e1 Mutated Mammalian Cells and
; TITLE OF INVENTION: Animals
; FILE REFERENCE: LEX-0102-USA
; CURRENT APPLICATION NUMBER: US/09/728,445
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358

```
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 799
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(407)
; OTHER INFORMATION: n = A,T,C or G
US-10-964-549-799

Alignment Scores:
Pred. No.: 1.36e-10 Length: 407
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-581-651D-6 (1-21) x US-09-728-445-799 (1-407)
Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysTyrGlyGlySer 20
Db 283 CAGCAGTGGGAACGGACCTACTAGGCAACGCCCTGTTGTACTGCTATGGAGGAAGC 342
Qy 21 Arg 21
Db 343 CGG 345

RESULT 2
US-10-964-549-799
; Sequence 799, Application US/10964549
; Publication No. US20050186677A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: Novel Mutated Mammalian Cells and Animals
; FILE REFERENCE: LEX-0286-USA
; CURRENT APPLICATION NUMBER: US/10/964,549
; CURRENT FILING DATE: 2004-10-13
; PRIOR APPLICATION NUMBER: US/09/750,456
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 09/728,445
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 799
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(407)
; OTHER INFORMATION: n = A,T,C or G
US-10-964-549-799

Alignment Scores:
Pred. No.: 1.36e-10 Length: 407
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-581-651D-6 (1-21) x US-10-964-549-799 (1-407)
Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysTyrGlyGlySer 20
Db 283 CAGCAGTGGGAACGGACCTACTAGGCAACGCCCTGTTGTACTGCTATGGAGGAAGC 342
Qy 21 Arg 21
Db 343 CGG 345

US-09-581-651D-6 (1-21) x US-09-925-302-105 (1-867)
Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysTyrGlyGlySer 20
Db 448 CAACAGTGGGAGCGGACCTACTAGGCAATGCGTTGGTTGTACTGTTATGGAGGAAGC 507
Qy 21 Arg 21
Db 508 CGA 510

RESULT 4
US-09-925-302-105
; Sequence 105, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-105

Alignment Scores:
Pred. No.: 3.29e-10 Length: 867
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-581-651D-6 (1-21) x US-09-925-302-105 (1-867)
Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysTyrGlyGlySer 20
Db 448 CAACAGTGGGAGCGGACCTACTAGGCAATGCGTTGGTTGTACTGTTATGGAGGAAGC 507
Qy 21 Arg 21
Db 508 CGA 510

RESULT 3
US-09-925-302-105
; Sequence 105, Application US/09925302
; Patent No. US2002004941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-105

Alignment Scores:
Pred. No.: 3.29e-10 Length: 867
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:49:16 ; Search time 21.1373 Seconds
(without alignments)

1625.652 Million cell updates/sec

Title: US-09-581-651D-6
Perfect score: 122
Sequence: 1 QWERTYLGALVCTCYGGR 21

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool_b/US09581651/runat_07112005_092224_28801/app_query.fasta_1.1834
-DB=Issued_Patents_NA -QPM=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09581651 @CGN 1 1 187 @runat_07112005_092224_28801 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122	100.0	8044	4	US-09-566-921-135
2	118	96.7	7679	4	US-09-220-132-38
3	118	96.7	7680	4	US-09-023-655-1289
4	118	96.7	7680	5	PCT-US95-09819-6
5	118	96.7	7705	1	US-08-259-569-16
6	118	96.7	7705	2	US-08-826-885-16
7	118	96.7	7705	6	US-08-826-885-16
8	118	96.7	7705	6	US-08-551-356-1
9	118	96.7	7803	2	US-08-551-356-1
10	118	96.7	7803	5	PCT-US93-12687-1
c 11	75	61.5	241	4	US-09-389-661-347
c 12	75	61.5	241	4	US-09-620-405B-347

c 13	75	61.5	241	4	US-09-433-826B-347	Sequence 347, App
c 14	75	61.5	241	4	US-09-604-287A-347	Sequence 347, App
c 15	75	61.5	241	4	US-09-834-759-347	Sequence 347, App
c 16	75	61.5	241	4	US-09-590-751A-347	Sequence 347, App
c 17	75	61.5	241	4	US-09-551-621-347	Sequence 347, App
c 18	75	61.5	247	3	US-09-397-787-118	Sequence 118, App
c 19	75	61.5	501	3	US-09-404-879A-252	Sequence 252, App
c 20	75	61.5	501	4	US-09-338-933-220	Sequence 252, App
c 21	75	61.5	501	4	US-09-215-681-252	Sequence 252, App
c 22	75	61.5	501	4	US-09-216-003A-252	Sequence 252, App
c 23	75	61.5	501	4	US-09-667-857-252	Sequence 252, App
c 24	75	61.5	567	1	US-08-142-449B-5	Sequence 5, Appli
c 25	75	61.5	828	3	US-09-404-879A-220	Sequence 220, App
c 26	75	61.5	828	4	US-09-338-933-220	Sequence 220, App
c 27	75	61.5	828	4	US-09-215-681-220	Sequence 220, App
c 28	75	61.5	828	4	US-09-216-003A-220	Sequence 220, App
c 29	75	61.5	828	4	US-09-667-857-220	Sequence 220, App
c 30	74	60.7	440	3	US-09-397-787-321	Sequence 321, App
c 31	63	51.6	68	6	5455158-4	Patent No. 5455158
c 32	63	51.6	68	6	5455158-4	Patent No. 5455158
c 33	55	45.1	57	1	US-08-259-569-21	Sequence 21, Appl
c 34	55	45.1	57	2	US-08-826-885-21	Sequence 21, Appl
c 35	55	45.1	57	6	5455158-5	Patent No. 5455158
c 36	55	45.1	57	6	5455158-5	Patent No. 5455158
c 37	55	45.1	70	1	US-08-259-569-22	Sequence 22, Appl
c 38	55	45.1	70	2	US-08-826-885-22	Sequence 22, Appl
c 39	54	44.3	69	1	US-08-259-569-19	Sequence 19, Appl
c 40	54	44.3	69	2	US-08-826-885-19	Sequence 19, Appl
c 41	53	43.4	56	1	US-08-259-569-20	Sequence 20, Appl
c 42	53	43.4	56	2	US-08-826-885-20	Sequence 20, Appl
c 43	52.5	43.0	986	1	US-07-637-250A-8	Sequence 8, Appli
c 44	52.5	43.0	986	1	US-08-145-061-8	Sequence 8, Appli
c 45	52.5	43.0	87563	3	US-09-453-702B-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1

US-09-566-921-135
; Sequence 135, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne P.
; APPLICANT: Tingley, Debora M.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 135
; LENGTH: 8044
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 427813.14
US-09-566-921-135

Alignment Scores:
Pred. No.: 3.48e-09
Score: 122.00
Matches: 21
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
Query Match: 100.00%
Indels: 0
DB: 4
Gaps: 0

US-09-581-651D-6 (1-21) x US-09-566-921-135 (1-8044)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 455 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTGGTTGTTGTTACTTGTATGGAGGAGC 514

Qy 21 Arg 21
Db 515 CGA 517

RESULT 2
US-09-220-132-38
; Sequence 38, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-38

Alignment Scores:
Pred. No.: 1.43e-08 Length: 7679
Score: 118.00 Matches: 20
Percent Similarity: 95.24% Conservative: 0
Best Local Similarity: 95.24% Mismatches: 1
Query Match: 96.72% Indels: 0
DB: 4 Gaps: 0

US-09-581-651D-6 (1-21) x US-09-220-132-38 (1-7679)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 107 CAACAGTGGAGCGGACCTACCTAGGTAAATGTGTGGTTGTACTTGTATTGAGGAGC 166

Qy 21 Arg 21
Db 167 CGA 169

RESULT 3
US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:

PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g31396
US-09-023-655-1289

Alignment Scores:
Pred. No.: 1.43e-08 Length: 7680
Score: 118.00 Matches: 20
Percent Similarity: 95.24% Conservative: 0
Best Local Similarity: 95.24% Mismatches: 1
Query Match: 96.72% Indels: 0
DB: 4 Gaps: 0

US-09-581-651D-6 (1-21) x US-09-023-655-1289 (1-7680)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 107 CAACAGTGGAGCGGACCTACCTAGGTAAATGTGTGGTTGTACTTGTATTGAGGAGC 166

Qy 21 Arg 21
Db 167 CGA 169

RESULT 4
PCT-US95-09819-6
; Sequence 6, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:30:56 ; Search time 68.8471 Seconds
(without alignments)
1805.663 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122

Sequence: 1 QQWERTYLGALVCTCYGSSR 21

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092222_28772/app_query.fasta_1.1834
-DB=N_Geneseq -QFT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651.QCGN 1.1 1052 @runat_07112005_092222_28772 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122	100.0	407	6	ABs69736
2	122	100.0	867	3	Abf18086 Novel mur
3	122	100.0	1929	13	Abf18086 Lung canc
4	122	100.0	1929	13	Adg17488 Nucleotid
5	122	100.0	2127	10	Adr97657 Human fib
					Add18477 Human pro

6	122	100.0	2147	2	AAX81299
7	122	100.0	2358	13	Adr67201 Human bla
8	122	100.0	2443	13	Adq38575 Human SNP
9	122	100.0	2488	13	Adq38581 Human SNP
10	122	100.0	3522	5	Aas86462 DNA encod
11	122	100.0	3669	12	Adp29134 Human sec
12	122	100.0	4295	8	Acc72037 BCU0770 g
13	122	100.0	6510	13	Adq38578 Human SNP
14	122	100.0	6816	10	Abx74443 Human CDN
15	122	100.0	7242	13	Adr66637 Human pro
16	122	100.0	7242	13	Adr65953 Human pro
17	122	100.0	7361	9	Adc06170 Human CDN
18	122	100.0	7550	8	Acc00412 Human cel
19	122	100.0	7677	13	Adn37820 Tumour-as
20	122	100.0	7795	10	Adj56196 Zabrifish
21	122	100.0	7823	13	Adq38582 Human SNP
22	122	100.0	7848	13	Adq38583 Human SNP
23	122	100.0	7867	4	Aai57803 Human pol
24	122	100.0	7935	13	Adq38577 Human SNP
25	122	100.0	7959	13	Adq38586 Human SNP
26	122	100.0	8013	13	Adq38579 Human SNP
27	122	100.0	8027	11	Adp64998 Human fib
28	122	100.0	8027	12	Adg89565 Human fib
29	122	100.0	8027	12	Adq29601 Human col
30	122	100.0	8027	13	Adr67200 Human bla
31	122	100.0	8039	3	Aac89889 Human FN
32	122	100.0	8044	6	Aas94866 Human DNA
33	122	100.0	8044	12	Adi61767 Human CDN
34	122	100.0	8062	4	Aai57802 Human pol
35	122	100.0	8066	4	Aai59591 Human pol
36	122	100.0	8066	4	Aai59589 Human pol
37	122	100.0	8066	4	Aai59588 Human pol
38	122	100.0	8066	4	Aai59590 Human pol
39	122	100.0	8137	4	Aai57805 Human pol
40	122	100.0	8155	13	Adq38584 Human SNP
41	122	100.0	8226	13	Adq38574 Human SNP
42	122	100.0	8230	4	Aai57804 Human pol
43	122	100.0	8278	13	Adq38587 Human SNP
44	122	100.0	8329	6	Abk63824 Rat seque
45	122	100.0	8329	10	Adb58331 Toxicity-

ALIGNMENTS

RESULT 1
ABS69736
ID ABS69736 standard; DNA; 407 BP.
XX
AC ABS69736;
XX
DT 21-NOV-2002 (first entry)
XX
DE Novel murine polynucleotide isolated using gene trap technology #799.
XX
KW Mouse; gene trapped sequence; GTS; functional genomic analysis;
KW phage display system; gene chip; temporal gene expression;
KW tissue specific gene expression; antisense inhibition; gene targeting;
KW development disorder; cell differentiation disorder; aging; cancer;
KW autoimmune disease; lupus; inflammatory disorder; skin disorder;
KW degenerative disorder; ds.
XX
OS Mus musculus.
XX
PN US2002102543-A1.
XX
PD 01-AUG-2002.
XX
PF 30-NOV-2000; 2000US-00728445.
XX
PR 01-DEC-1999; 99US-0168358P.
XX
FA (FRIE/) FRIEDRICH G.
FA (ZAMB/) ZAMBROWICZ B.

PA (SAND/) SANDS A T.
XX Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2002-690598/74.
XX Novel murine polynucleotides that individually identify novel genes into
DR which a retroviral gene trap vector has integrated, useful in genomic
PT analysis and in discovery, development of therapeutic and diagnostic
PT agents.
XX Claim 1; Page 264; 296pp; English.
PS The invention describes an isolated murine polynucleotide (I) comprising
XX a contiguous stretch of at least 60 nucleotides of one of 265-677
CC nucleotide 891 ONNIBANK gene trapped sequences (GTSs) (S), given in the
CC specification. The novel genes and cells are useful in functional genomic
CC analysis and in the discovery and development of new therapeutic and
CC diagnostic agents and methods. (I) is useful for identifying the coding
CC regions of the murine genome, to isolate cDNAs, genomic clones, or full-
CC length genes/polynucleotides or homologues, heterologues, paralogues, or
CC orthologues that are capable of hybridizing to one or more of the GTSs
CC under stringent conditions. (I) can be incorporated into a phage display
CC system that can be used to screen for proteins, or other ligands, that
CC are capable of binding an amino acid sequence encoded by an
CC oligonucleotide or polynucleotide sequence in at least one of the TS
CC sequences. (I) is useful in addressable arrays, such as gene chips, to
CC identify and characterize temporal and tissue specific gene expression,
CC to identify the gene of interest from many sources and for genetic
CC manipulations such as antisense inhibition and gene targeting. Decreasing
CC the level of expression of (I) and/or down regulating the activity of
CC peptides or proteins encoded by (I) is useful for treating development
CC and cell differentiation disorders, aging, cancer, autoimmune disease,
CC lupus, inflammatory disorders, skin disorders and degenerative disorders.
CC This sequence represents a murine cDNA isolated using gene trap
XX technology
SQ Sequence 407 BP; 82 A; 114 C; 131 G; 74 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 4.92e-10 Length: 407
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-6 (1-21) x ABS69736 (1-407)

Qy 1 GlnGlnTrpGluArgThrTyrrleuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 283 CAGCAGTGGGACGCGACCTACCTAGGCAACGCCCTGGTTGTACCTGCTATGGAGGAAGC 342

Qy 21 Arg 21
Db 343 CGG 345

RESULT 2
AAFI8086
ID AAFI8086 standard; DNA; 867 BP.
XX AAFI8086;
AC AAFI8086;
XX 14-MAR-2001 (first entry)
DT Lung cancer associated polynucleotide sequence SEQ ID 105.
DE Human; lung cancer associated protein; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active; vulnerary;
KW gastrointestinal; nephrotropic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease; ds.
XX

OS Homo sapiens.
XX WO200055180-A2.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US005918.
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX Ruben SM;
XX WPI; 2000-587514/55.
XX P-PSDS; AAB58210.
XX Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer.
XX Claim 1; Page 576-577; 1425pp; English.
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective, cytostatic, cardioactive, and
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the protein
CC or polynucleotide sequences. The lung cancer associated polynucleotide
CC sequences may be used for detection of lung cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The proteins may be used to treat disorders such as
CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
CC cardiovascular, renal, and proliferative disorders. The proteins may also
CC be used in the treatment of wounds and infectious diseases.
CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
CC used in the course of the invention for the identification and
CC characterisation of the polynucleotide and protein sequences
XX
SQ Sequence 867 BP; 195 A; 223 C; 268 G; 176 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.: 1.22e-09 Length: 867
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-581-651D-6 (1-21) x AAF18086 (1-867)

Qy 1 GlnGlnTrpGluArgThrTyrrleuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 448 CAACAGTGGGACGCGACCTACCTAGGCAATGCGTTGGTTTGTACTTGTATGGAGGAAGC 507

Qy 21 Arg 21
Db 508 CGA 510

RESULT 3
ADSI17488
ID ADSI17488 standard; DNA; 1929 BP.
XX ADSI17488;
AC ADSI17488;
XX 02-DEC-2004 (first entry)
DT Nucleotide sequence of human fibronectin 1.
DE cell state; time-lapse profile; protein-protein interaction;
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 562.251 Seconds
(without alignments)
1809.796 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122
Sequence: 1 QQMERTYLGALVCTCYGGR 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092223_28778/app_query.fasta_1.1834
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CGN 1.1 9936 @runat_07112005_092223_28778 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_btg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ats:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	122	100.0	706 6 CQ731570 Sequence
2	122	100.0	1929 6 CQ871810 Sequence
3	122	100.0	1929 6 CQ871828 Sequence
4	122	100.0	2147 6 BD137021 Polypepti

5	122	100.0	2147	6	AX003229	Sequence
6	122	100.0	2147	9	HA276395	Homo sapi
7	122	100.0	2192	9	HA535086	Homo sapi
8	122	100.0	2358	6	CQ875358	Sequence
9	122	100.0	3059	10	BC051082	Mus muscu
10	122	100.0	3796	10	RNF1B1	X05831 Rat fibro
11	122	100.0	4321	9	HSB06214	Sequence
12	122	100.0	6816	6	CQ715726	Sequence
13	122	100.0	7502	9	HSB04082	Sequence
14	122	100.0	7544	9	HSB03509	Homo sapi
15	122	100.0	7753	9	AB191261	Homo sapi
16	122	100.0	7868	9	HSB06653	Sequence
17	122	100.0	7951	9	HSB06901	Homo sapi
18	122	100.0	7951	9	HSB06902	Homo sapi
19	122	100.0	8027	6	CQ833991	Sequence
20	122	100.0	8027	6	CQ875357	Sequence
21	122	100.0	8030	9	HSB06170	Homo sapi
22	122	100.0	8035	9	HSB06171	Homo sapi
23	122	100.0	8042	9	HSB06805	Homo sapi
24	122	100.0	8044	6	AR454662	Sequence
25	122	100.0	8044	6	AX281712	Sequence
26	122	100.0	8320	9	HSB06267	Homo sapi
27	122	100.0	8329	6	AX402055	Sequence
28	122	100.0	8329	10	RNF1B1	Sequence
29	122	100.0	8411	9	HSB06992	Homo sapi
30	122	100.0	8421	9	HSB06903	Homo sapi
31	122	100.0	143947	9	AC012462	Homo sapi
32	122	100.0	185651	2	AC124821	Mus muscu
33	122	100.0	197121	2	AC127101	Rattus no
34	122	100.0	204737	2	AC091456	Mus muscu
35	122	100.0	255929	2	AC095457	Rattus no
36	118	96.7	7679	6	AR274901	Sequence
37	118	96.7	7680	6	AR380744	Sequence
38	118	96.7	7680	6	AX277596	Sequence
39	118	96.7	7680	6	AX335368	Sequence
40	118	96.7	7680	9	HSF1B1	Human mRNA
41	118	96.7	7705	6	AI4133	Fibronectin
42	118	96.7	7705	6	AR034630	Sequence
43	118	96.7	7705	6	E01162	CDNA encodi
44	118	96.7	7705	6	I70110	Sequence 16
45	118	96.7	7705	6	AR364992	Sequence

ALIGNMENTS

RESULT 1	CQ731570	706 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	Sequence 17504 from Patent WO02068579.				
DEFINITION	CQ731570				
ACCESSION	CQ731570.1	GI:42308500			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.				
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof				
JOURNAL	Patent: WO 02068579-A 17504 06-SEP-2002; PE Corporation (NY) (US)				
FEATURES	Location/Qualifiers				
source	1..706				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN

Alignment Scores: 1.15e-09 Length: 706
Pred. No.: 122.00 Matches: 21
Score:

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-6 (1-21) x CQ731570 (1-706)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 187 CAACAGTGGGAGCGACCTACCTAGCAATCGTTGGTTTGTACTTGTATTGAGGAGC 246
Qy 21 Arg 21
Db 247 CGA 249

RESULT 2
LOCUS CQ871810 1929 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 1 from Patent WO2004079332.
ACCESSION CQ871810
VERSION CQ871810.1 GI:52745842
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Miyake, M.A., Yoshikawa, T.A., Uchimura, E.A. and Miyake, J.A.
TITLE Composition and method for increasing efficiency of introduction of
JOURNAL target substance into cell
Patent: WO 2004079332-A 1 16-SEP-2004;
National Institute of Advanced Industrial Science and Tec hnology
(JP)
FEATURES
source Location/Qualifiers
1..1929
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..1929
/note="unnamed protein product; fibronectin 1"
CDS
ORIGIN
Alignment Scores: 2.78e-09 Length: 1929
Pred. No.: 122.00 Matches: 21
Score: 122.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

US-09-581-651D-6 (1-21) x CQ871810 (1-1929)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 187 CAACAGTGGGAGCGACCTACCTAGCAATCGTTGGTTTGTACTTGTATTGAGGAGC 246
Qy 21 Arg 21
Db 247 CGA 249

RESULT 4
LOCUS BD137021 2147 bp DNA linear PAT 18-SEP-2002
DEFINITION Polypeptides, polynucleotides and uses thereof.
ACCESSION BD137021
VERSION BD137021.1 GI:23231966
KEYWORDS JP 2002508179-A/1.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2147)
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 6.14379 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-5
Perfect score: 110
Sequence: 1 ISKYLWRPVSIPPNLGY 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	749	2 Q7Z391	Q7Z391 homo sapien
2	61	55.5	2265	1 F1NC_BOVIN	P07589 bos taurus
3	56	50.9	1103	2 Q6MZFA	Q6MZFA homo sapien
4	56	50.9	2193	2 Q6MZM7	Q6MZM7 homo sapien
5	56	50.9	2240	2 Q6MDP8	Q6MDP8 homo sapien
6	56	50.9	2267	2 Q6MDP9	Q6MDP9 homo sapien
7	56	50.9	2296	2 Q6N0A6	Q6N0A6 homo sapien
8	56	50.9	2357	2 Q68D74	Q68D74 homo sapien
9	56	50.9	2386	1 F1NC_HUMAN	P02751 homo sapien
10	56	50.9	2444	2 Q6N025	Q6N025 homo sapien
11	56	50.9	2477	2 Q6MZU5	Q6MZU5 homo sapien
12	53	48.2	1367	1 IG1R_HUMAN	P08069 homo sapien
13	53	48.2	2477	1 F1NC_MOUSE	P11276 mus musculus
14	53	48.2	2477	1 F1NC_RAT	P04937 rattus norv
15	52	47.3	1358	2 Q73798	Q73798 xenopus lae
16	51	46.4	196	2 Q57542	Q57542 oenococcus
17	51	46.4	1253	2 Q57566	Q57566 canis famil
18	51	46.4	2936	2 QYRKR8	QYRKR8 canis famil
19	50	45.5	275	2 Q825H3	Q825H3 streptomyce
20	50	45.5	293	2 Q9HTT0	Q9HTT0 pseudomonas
21	50	45.5	485	2 Q8H633	Q8H633 oryza sativ
22	50	45.5	1370	1 IG1R_RAT	P24062 rattus norv
23	50	45.5	1371	2 Q9QVW4	Q9QVW4 rattus sp.
24	50	45.5	1373	1 IG1R_MOUSE	Q60751 mus musculu
25	49.5	45.0	372	2 Q9HG66	Q9HG66 saitoella c
26	48	43.6	189	2 Q8BBK5	Q8BBK5 hepatitis c
27	48	43.6	542	2 Q8Y2T1	Q8Y2T1 raietonia s
28	48	43.6	2515	2 Q77365	Q77365 plasmodium s
29	47.5	43.2	922	2 Q93405	Q93405 brachydanio
30	47.5	43.2	2478	2 Q93406	Q93406 brachydanio
31	47	42.7	126	1 PBE3_PHOLL	Q7N3R1 photorhabdu

32 47 42.7 260 2 Q88CN7
33 47 42.7 283 2 Q6H4G2
34 47 42.7 439 1 HGD_DROME
35 47 42.7 645 2 Q8PUN5
36 46.5 42.3 241 2 Q6SHD4
37 46 41.8 87 2 Q8H444
38 46 41.8 196 1 K18A_MOUSE
39 46 41.8 196 2 Q6PDU2
40 46 41.8 239 2 Q8S2P6
41 46 41.8 471 2 Q8P3X6
42 46 41.8 471 2 Q8PFG0
43 46 41.8 539 2 Q84J36
44 46 41.8 610 2 Q8A6W7
45 46 41.8 1407 2 Q6SK65

Q88cn7 pseudomonas
Q6h4g2 oryza sativ
Q9vkj0 drosophila
Q8pun5 methanosarc
Q6shd4 uncultured
Q8h444 oryza sativ
Q9n85 mus musculus
Q6pdu2 rattus norv
Q8s2p6 oryza sativ
Q8p3x6 xanthomonas
Q8pfg0 xanthomonas
Q84j36 oryza sativ
Q8a6w7 bacteroides
Q6sk65 arthrobacte

ALIGNMENTS

RESULT 1
Q7Z391
ID Q7Z391 PRELIMINARY; PRT; 749 AA.
AC Q7Z391;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloecher H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -.
DR HSSP; Q96KP7; 1PBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctn.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 100.0%; Score 110; DB 2; Length 749;

Best Local Similarity 100.0%; Pred.No. 2.2e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLWRPVSIPPNLGY 20

Db 730 ISKYLWRPVSIPPNLGY 749

RESULT 2

F1NC_BOVIN

ID F1NC_BOVIN STANDARD; PRT; 2265 AA.

AC P07589;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Fibronectin (FN).

GN Name=FN1;

OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovinae; Bos.
 OK NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=87054047; PubMed=3780752;
 RX Skorstengaard K., Jensen M.S., Sahl P., Petersen T.E., Magnusson S.;
 RT "Complete primary structure of bovine plasma fibronectin.";
 RL Eur. J. Biochem. 161:441-453(1986).
 RN [2]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=83117805; PubMed=6218503;
 RA Petersen T.E., Thorgersen H.C., Skorstengaard K., Vibe-Pedersen K.,
 SA Sahl P., Sottrup-Jensen L., Magnusson S.;
 RT "Partial primary structure of bovine plasma fibronectin: three types
 of internal homology.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:137-141(1983).
 RN [3]
 RP SEQUENCE OF 2170-2265 FROM N.A.
 RX MEDLINE=83221567; PubMed=6304699;
 RA Kornblith A.R., Vibe-Pedersen K., Baralle F.E.;
 RT "Isolation and characterization of cDNA clones for human and bovine
 fibronectins";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3218-3222(1983).
 CC -|- FUNCTION: Fibronectins bind cell surfaces and various compounds
 including collagen, fibrin, heparin, DNA, and actin. Fibronectins
 are involved in cell adhesion, cell motility, opsonization, wound
 healing, and maintenance of cell shape.
 CC -|- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced
 variants, connected by 2 disulfide bonds near the carboxyl ends;
 to a lesser extent homodimers.
 CC -|- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 CC -|- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=1;
 Comment=A number of isoforms are produced. Each of the "extra
 domain" and the connecting strand 3 are present in some forms of
 fibronectin and absent in others;
 Name=1;
 IsoId=P07589-1; Sequence=Displayed;
 CC -|- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
 by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
 forms), made by fibroblasts, epithelial and other cell types, is
 deposited as fibrils in the extracellular matrix.
 CC -|- PTM: Sulfated (By similarity).
 CC -|- SIMILARITY: Contains 12 fibronectin type I domains.
 CC -|- SIMILARITY: Contains 2 fibronectin type II domains.
 CC -|- SIMILARITY: Contains 15 fibronectin type III domains.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; K00800; AAA30521.2; .
 DR PIR; A26452; FNBO.
 DR HSP; P08253; 1KSO.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrctnl.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR008957; FN III-like.
 DR InterPro; IPR000562; FN Type II.
 DR InterPro; IPR003962; FnIII_subd.
 DR Pfam; PF00039; fn1; 12.
 DR Pfam; PF00040; fn2; 2.
 DR Pfam; PF00041; fn3; 15.
 DR PRINTS; PR00012; FNTYPEI.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00014; FNTYPEIII.

DR ProDom: PD000995; FN Type II; 2.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01253; FIBRONECTIN_1; 12.
 DR PROSITE; PS00023; FIBRONECTIN_2; 2.
 DR PROSITE; PS00853; FN3; 15.
 KW Acute phase; Alternative splicing; Cell adhesion; Cell shape;
 KW Direct protein sequencing; Glycoprotein; Heparin-binding;
 KW Phosphorylation; Plasma; Pyrrolidone carboxylic acid; Repeat;
 KW Sulfation.
 FT MOD_RES 1 1 Pyrrolidone carboxylic acid.
 FT DOMAIN 21 241 Fibrin- and heparin-binding 1.
 FT DOMAIN 277 577 Collagen-binding.
 FT DNA_BIND 876 1141 Cell-attachment.
 FT DOMAIN 1236 1509 Heparin-binding 2.
 FT DOMAIN 1600 1870 Fibrin-binding 2.
 FT DOMAIN 1991 2216 Fibronectin type-I 1.
 FT DOMAIN 19 59 Fibronectin type-I 2.
 FT DOMAIN 64 107 Fibronectin type-I 3.
 FT DOMAIN 108 151 Fibronectin type-I 4.
 FT DOMAIN 153 197 Fibronectin type-I 5.
 FT DOMAIN 198 242 Fibronectin type-I 6.
 FT DOMAIN 275 314 Fibronectin type-I 1.
 FT DOMAIN 314 373 Fibronectin type-II 1.
 FT DOMAIN 374 438 Fibronectin type-II 2.
 FT DOMAIN 437 480 Fibronectin type-I 7.
 FT DOMAIN 485 527 Fibronectin type-I 8.
 FT DOMAIN 528 571 Fibronectin type-I 9.
 FT DOMAIN 576 668 Fibronectin type-III 1.
 FT DOMAIN 689 778 Fibronectin type-III 2.
 FT DOMAIN 780 867 Fibronectin type-III 3.
 FT DOMAIN 877 964 Fibronectin type-III 4.
 FT DOMAIN 965 1053 Fibronectin type-III 5.
 FT DOMAIN 1056 1141 Fibronectin type-III 6.
 FT DOMAIN 1142 1234 Fibronectin type-III 7.
 FT DOMAIN 1235 1325 Fibronectin type-III 8.
 FT DOMAIN 1326 1415 Fibronectin type-III 9.
 FT DOMAIN 1416 1505 Fibronectin type-III 10.
 FT DOMAIN 1510 1599 Fibronectin type-III 11 (extra domain).
 FT DOMAIN 1602 1689 Fibronectin type-III 12.
 FT DOMAIN 1692 1780 Fibronectin type-III 13.
 FT DOMAIN 1781 1870 Fibronectin type-III 14.
 FT DOMAIN 1871 1990 Connecting strand 3 (CS-3) (V region).
 FT DOMAIN 1979 2069 Fibronectin type-III 15.
 FT DOMAIN 2083 2127 Fibronectin type-I 10.
 FT DOMAIN 2128 2170 Fibronectin type-I 11.
 FT DOMAIN 2172 2215 Fibronectin type-I 12.
 FT SITE 1493 1495 Cell attachment site.
 FT DISULFID 21 47
 FT DISULFID 45 56
 FT DISULFID 66 94
 FT DISULFID 92 104
 FT DISULFID 110 138
 FT DISULFID 136 148
 FT DISULFID 155 184
 FT DISULFID 182 194
 FT DISULFID 200 229
 FT DISULFID 227 239
 FT DISULFID 277 304
 FT DISULFID 302 311
 FT DISULFID 329 355
 FT DISULFID 343 370
 FT DISULFID 389 415
 FT DISULFID 403 430
 FT DISULFID 439 467
 FT DISULFID 465 477
 FT DISULFID 487 514
 FT DISULFID 512 524
 FT DISULFID 530 558
 FT DISULFID 556 568
 FT DISULFID 2085 2114
 FT DISULFID 2112 2124
 FT DISULFID 2130 2157
 FT DISULFID 2155 2167

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 1.51634 Seconds
(without alignments)
1269.066 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKYLKRPVSPRNLGY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	55.5	2265	1 FNBO	fibronectin - bovi
2	56	50.9	2386	1 FNHU	fibronectin precu
3	53	48.2	1367	1 IGHUT1	insulin-like growt
4	53	48.2	2477	2 S14428	fibronectin precu
5	51	46.4	196	2 S42038	hypothetical prote
6	50	45.5	293	2 B82988	hypothetical prote
7	50	45.5	1371	2 A33837	insulin-like growt
8	48	43.6	2523	2 T18477	hypothetical prote
9	45	40.9	203	2 A75453	conserved hypothet
10	45	40.9	224	2 D82966	conserved hypothet
11	44	40.0	196	2 A69193	hypothetical prote
12	44	40.0	210	2 I55609	insulin receptor p
13	44	40.0	719	2 T09543	deoxyxylulose synt
14	44	40.0	958	2 T04870	cellulose synthase
15	44	40.0	1372	2 A34157	insulin receptor p
16	44	40.0	1382	1 INHUR	insulin receptor p
17	44	40.0	1383	2 A36080	insulin receptor p
18	44	40.0	2481	2 A43908	fibronectin - Afri
19	44	40.0	2944	2 A54849	collagen alpha 1(V
20	43.5	39.5	429	2 D82669	citrate synthase x
21	43.5	39.5	442	2 S20908	tubulin beta chain
22	43.5	39.5	447	2 J00423	tubulin beta chain
23	43.5	39.5	447	2 A25377	tubulin beta chain
24	43	39.1	219	2 T46472	hypothetical prote
25	43	39.1	321	2 H70669	hypothetical prote
26	43	39.1	460	2 D96813	hypothetical prote
27	43	39.1	512	1 NTAIWB	nitrogenase (EC 1.
28	43	39.1	512	2 A51986	nitrogenase molybd
29	43	39.1	561	2 B90414	hypothetical prote

RESULT 1

FNBO

fibronectin - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 31-Dec-1988 #sequence, revision 31-Dec-1988 #text_change 09-Jul-2004

C:Accession: A26452; B21165; A23292

R:Skorsteengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.

Rur: J. Blochem. 161, 441-453, 1986

A:Title: Complete primary structure of bovine plasma fibronectin.

A:Reference number: A26452; MUID:87054047; PMID:3780752

A:Accession: A26452

A:Molecule type: protein

A:Residues: 1-2265 <SKO>

A:Cross-references: UNIPROT:P07589

R:Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.

Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983

A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins

A:Reference number: A21165; MUID:83221567; PMID:6304699

A:Accession: B21165

A:Molecule type: mRNA

A:Residues: 2170-2265 <KOR>

A:Cross-references: GB:K00800; NID:6163055; PIDN:AAA30521.2; PID:95713323

R:Petersen, T.E.; Thøgersen, H.C.; Skorsteengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sottr

Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983

A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal

A:Reference number: A23292; MUID:83117805; PMID:6218503

A:Accession: A23292

A:Molecule type: protein

A:Residues: 1-16, 'C', 18-20, 'S', 22-432; 447-463; 1367-1517; 1567-1673; 2062-2176, 'N', 2178-226

C:Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.

C:Comment: The plasma fibronectin molecule consists of two chains, which are connected b

C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib

aling, and maintenance of cell shape.

C:Comment: Plasma fibronectin is synthesized by hepatocytes.

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe

C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu

F:21-241/Domain: fibrin and heparin binding <FBR>

F:21-56/Domain: fibronectin type I repeat homology <1F1>

F:66-104/Domain: fibronectin type I repeat homology <1F2>

F:110-148/Domain: fibronectin type I repeat homology <1F3>

F:155-194/Domain: fibronectin type I repeat homology <1F4>

F:200-239/Domain: fibronectin type I repeat homology <1F5>

F:277-577/Domain: collagen binding <CBR>

F:277-311/Domain: fibronectin type I repeat homology <1F6>

F:329-370/Domain: fibronectin type II repeat homology <2F1>

F:389-430/Domain: fibronectin type II repeat homology <2F2>

F:439-477/Domain: fibronectin type I repeat homology <1F7>

F:487-524/Domain: fibronectin type I repeat homology <1F8>

F:530-568/Domain: fibronectin type I repeat homology <1F9>

F:578-661/Domain: fibronectin type III repeat homology <FN3A>

F:688-770/Domain: fibronectin type III repeat homology <FN3B>

F:779-860/Domain: fibronectin type III repeat homology <FN3C>

F:779-860/Domain: fibronectin type III repeat homology <FN3C>

F:779-860/Domain: fibronectin type III repeat homology <FN3C>

F:779-860/Domain: fibronectin type III repeat homology <FN3C>

F:875-957/Domain: fibronectin type III repeat homology <FN3D>
F:955-1046/Domain: fibronectin type III repeat homology <FN3E>
F:1055-1134/Domain: fibronectin type III repeat homology <FN3F>
F:1142-1227/Domain: fibronectin type III repeat homology <FN3G>
F:1235-1318/Domain: fibronectin type III repeat homology <FN3H>
F:1326-1404/Domain: fibronectin type III repeat homology <GN3I>
F:1410-1517/Domain: cell attachment <CAD>
F:1416-1502/Domain: fibronectin type III repeat homology <FN3J>
F:1493-1495/Region: cell attachment (R-G-D) motif
F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>
F:1600-1870/Domain: heparin binding <HB2>
F:1600-1682/Domain: fibronectin type III repeat homology <FN3L>
F:1692-1773/Domain: fibronectin type III repeat homology <FN3M>
F:1781-1863/Domain: fibronectin type III repeat homology <FN3N>
F:1970-1972/Region: cell attachment (R-G-D) motif
F:1982-2062/Domain: fibronectin type III repeat homology <FN3O>
F:1985-2216/Domain: fibrin binding <FB2>
F:2085-2124/Domain: fibronectin type I repeat homology <1F10>
F:2130-2167/Domain: fibronectin type I repeat homology <1F11>
F:2174-2209/Domain: fibronectin type I repeat homology <1F12>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Cross-link: isopeptide (Gln) (interchain to lys N6-amino of fibrin) #status expirime
F:21-47,45-56,66-94,92-104,110-138,136-148,155-184,182-194,200-229,227-239,277-304,302-3
7,2155-2167,2174-2200,2198-2209/Diulfide bonds: #status predicted
F:399,497,511,846,976,1213,1987/Binding site: carbohydrate (Asn) (covalent) #status expe
F:1205,1692/Binding site: carbohydrate (Asn) (covalent) #status absent
F:1943,1944/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:12246/bisulfide bonds: interchain (to 2250) #status predicted
F:2250/bisulfide bonds: interchain (to 2246) #status predicted
F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 55.5%; Score 61; DB 1; Length 2265;
Best Local Similarity 68.1%; Pred. No. 0.78;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ISKYLWRKPSIPPR 16
|||||||:|:|
DB 607 ISKYLWRKPNSPDR 522

RESULT 2
FNHU
N/Alternate names: fibronectin [validated] - human
C/Species: Homo sapiens (man)
C/Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004
C/Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22
R/Dean, D.C.; Bowlus, C.L.; Bourgeois, S.
Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
A/Title: Cloning and analysis of the promoter region of the human fibronectin gene.
A/Reference number: A26460; MUID:87175578; PMID:3031656
A/Accession: A26460
A/Molecule type: DNA
A/Residues: 1-49 <DEA>
A/Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAA5337
R/Oldberg, A.; Ruoslahti, E.
J. Biol. Chem. 261, 2113-2116, 1986
A/Title: Evolution of the fibronectin gene.
A/Reference number: A26284; MUID:86111901; PMID:3003095
A/Accession: A26284
A/Molecule type: DNA
A/Residues: 1447-1540 <OLD>
A/Cross-references: GB:M12549; NID:g182688
A/Note: the authors translated the codon TTC for residue 1494 as Glu
R/Paoletti, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.
Nucleic Acids Res. 16, 3545-3557, 1988
A/Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B
A/Reference number: S00848; MUID:88233940; PMID:3375063
A/Accession: S03917
A/Molecule type: DNA
A/Residues: 1594-1767, 'V', 1769-1783 <PAO>
A/Cross-references: EMBL:X07718; NID:g31402
A/Note: the authors translated the codon AAC for residue 1631 as Asp

R/Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
FEBS Lett. 207, 287-291, 1986
A/Title: Donor and acceptor splice signals within an exon of the human fibronectin gene:
A/Reference number: A24854; MUID:87030929; PMID:3770201
A/Accession: A24854
A/Molecule type: DNA
A/Residues: 1992-2147 <VIB>
A/Cross-references: GB:X04530; NID:g31436
R/Gutman, A.; Yamada, K.M.; Kornblihtt, A.
FEBS Lett. 207, 145-148, 1986
A/Title: Human fibronectin is synthesized as a pre-propolypeptide.
A/Reference number: A24476; MUID:87030890; PMID:3770189
A/Accession: A24476
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-14, 'Q', 16-38 <GUT>
R/Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
EMBO J. 4, 1755-1759, 1985
A/Title: Primary structure of human fibronectin: differential splicing may generate at l
A/Reference number: A91008; MUID:85284965; PMID:2992939
A/Accession: A91008
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 32-1344, 1346-2080; 2112-2386 <KOR>
R/Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A/Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
A/Reference number: A93529; MUID:84272258; PMID:6462919
A/Accession: A93529
A/Molecule type: mRNA
A/Residues: 973-2080; 2112-2386 <KO2>
A/Cross-references: GB:X00739
R/Oldberg, A.; Limney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A/Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
A/Reference number: A21011; MUID:83290929; PMID:6688418
A/Accession: A21011
A/Molecule type: mRNA
A/Residues: 1434-1537 <OL2>
A/Cross-references: GB:K00055; NID:g182680; PIDN:AAA52459.1; PID:g182683
R/Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A/Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
A/Reference number: A90495; MUID:85280409; PMID:2992573
A/Accession: A90495
A/Molecule type: mRNA
A/Residues: 1594-2386 <BER>
R/Umezawa, K.; Kornblihtt, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A/Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A/Reference number: A22245; MUID:85231203; PMID:2989004
A/Accession: A22245
A/Molecule type: mRNA
A/Residues: 1948-2067 <UME>
A/Cross-references: GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:g182706
A/Accession: B22245
A/Molecule type: mRNA
A/Residues: 1975-1991; 2017-2039 <UM2>
A/Cross-references: GB:M27590
R/Sekiguchi, K.; Klog, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A/Title: Human liver fibronectin complementary DNAs: identification of two different mes
A/Reference number: I52394; MUID:87026578; PMID:3021206
A/Accession: I65273
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
A/Cross-references: GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:g182704
R/Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A/Title: Isolation and characterization of cDNA clones for human and bovine fibronectin:

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 6.77124 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKYLWRPVSIPPRNLGY 20

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	642	16	US-10-741-601-354
2	110	100.0	642	17	US-10-741-600-1066
3	110	100.0	657	16	US-10-741-601-359
4	110	100.0	657	17	US-10-741-600-1072
5	56	50.9	26	13	US-10-156-820-2
6	56	50.9	26	13	US-10-156-820-78
7	56	50.9	75	15	US-10-005-171-1
8	56	50.9	75	15	US-10-431-642-1
9	56	50.9	984	16	US-10-741-601-356
10	56	50.9	984	17	US-10-741-600-1069
11	56	50.9	1173	18	US-10-450-763-52634

12	56	50.9	2220	15	US-10-236-392-4	Sequence 4, Appli
13	56	50.9	2296	16	US-10-741-601-363	Sequence 363, App
14	56	50.9	2296	17	US-10-741-600-1075	Sequence 1075, Ap
15	56	50.9	2320	14	US-10-279-733-8	Sequence 8, Appli
16	56	50.9	2320	15	US-10-236-392-2	Sequence 2, Appli
17	56	50.9	2328	14	US-10-171-311-64	Sequence 64, Appli
18	56	50.9	2328	15	US-10-236-031B-70	Sequence 70, Appli
19	56	50.9	2328	15	US-10-374-979-98	Sequence 98, Appli
20	56	50.9	2328	15	US-10-182-936A-98	Sequence 98, Appli
21	56	50.9	2328	16	US-10-477-238A-677	Sequence 677, App
22	56	50.9	2328	16	US-10-680-287A-677	Sequence 677, App
23	56	50.9	2328	17	US-10-477-173-677	Sequence 677, App
24	56	50.9	2355	15	US-10-144-194A-104	Sequence 104, App
25	56	50.9	2355	15	US-10-360-101-235	Sequence 235, App
26	56	50.9	2355	15	US-10-447-161-3	Sequence 3, Appli
27	56	50.9	2355	16	US-10-734-564-94	Sequence 94, Appli
28	56	50.9	2355	16	US-10-741-601-357	Sequence 357, App
29	56	50.9	2355	16	US-10-741-601-366	Sequence 366, App
30	56	50.9	2355	16	US-10-491-566-104	Sequence 104, App
31	56	50.9	2355	17	US-10-741-600-1067	Sequence 1067, Ap
32	56	50.9	2355	17	US-10-741-600-1078	Sequence 1078, Ap
33	56	50.9	2355	17	US-10-852-335A-147	Sequence 147, App
34	56	50.9	2355	18	US-10-287-436A-436	Sequence 436, App
35	56	50.9	2355	18	US-10-287-436A-1137	Sequence 1137, Ap
36	56	50.9	2355	20	US-11-040-130-28	Sequence 28, Appli
37	56	50.9	2386	10	US-09-561-403-1	Sequence 1, Appli
38	56	50.9	2386	16	US-10-741-601-360	Sequence 360, App
39	56	50.9	2386	16	US-10-788-792-206	Sequence 206, App
40	56	50.9	2386	16	US-10-618-281-32	Sequence 32, Appli
41	56	50.9	2386	17	US-10-741-600-1071	Sequence 1071, Ap
42	56	50.9	2386	17	US-10-868-577A-59	Sequence 59, Appli
43	56	50.9	2386	18	US-10-485-758-4	Sequence 4, Appli
44	56	50.9	2386	18	US-10-485-758-9	Sequence 9, Appli
45	56	50.9	2474	18	US-10-450-763-52638	Sequence 52638, A

ALIGNMENTS

RESULT 1
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match 100.0%; Score 110; DB 16; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ISKYLWRPVSIPPRNLGY 20
Db 623 ISKYLWRPVSIPPRNLGY 642

RESULT 2
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

```
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match 100.0%; Score 110; DB 17; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLRWRPVSIPPRNLGY 20
Db 623 ISKYLRWRPVSIPPRNLGY 642

RESULT 3
US-10-741-601-359
; Sequence 359, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-359

Query Match 100.0%; Score 110; DB 16; Length 657;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLRWRPVSIPPRNLGY 20
Db 638 ISKYLRWRPVSIPPRNLGY 657

RESULT 4
US-10-741-600-1072
; Sequence 1072, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1072
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1072

Query Match 100.0%; Score 110; DB 17; Length 657;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ISKYLRWRPVSIPPRNLGY 20
Db 638 ISKYLRWRPVSIPPRNLGY 657

RESULT 5
US-10-156-820-2
; Sequence 2, Application US/10156820
; Publication No. US20020150558A1
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/10/156,820
; CURRENT FILING DATE: 2002-06-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-156-820-2

Query Match 50.9%; Score 56; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLRWRP 10
Db 15 ISKYLRWRP 24

RESULT 6
US-10-156-820-78
; Sequence 78, Application US/10156820
; Publication No. US20020150558A1
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/10/156,820
; CURRENT FILING DATE: 2002-06-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotome
US-10-156-820-78

Query Match 50.9%; Score 56; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLRWRP 10
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 2.11765 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKYLWRPVSIPRNLGY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	50.9	26	4	US-09-155-613A-2
2	56	50.9	26	4	US-09-155-613A-78
3	56	50.9	31	1	US-07-829-462-2
4	56	50.9	31	1	US-08-340-812-2
5	56	50.9	31	1	US-08-459-064B-2
6	56	50.9	31	2	US-08-460-421A-2
7	56	50.9	31	2	US-08-717-169-4
8	56	50.9	31	4	US-09-228-901A-4
9	56	50.9	31	5	PCT-US93-00909-2
10	56	50.9	56	2	US-08-717-169-6
11	56	50.9	56	4	US-09-228-901A-6
12	56	50.9	66	1	US-08-340-812-12
13	56	50.9	66	1	US-08-459-064B-12
14	56	50.9	66	2	US-08-460-421A-12
15	56	50.9	75	2	US-08-717-169-3
16	56	50.9	75	4	US-09-228-901A-3
17	56	50.9	85	1	US-08-340-812-11
18	56	50.9	85	1	US-08-459-064B-11
19	56	50.9	85	2	US-08-460-421A-11
20	56	50.9	89	1	US-08-340-812-13
21	56	50.9	89	1	US-08-459-064B-13
22	56	50.9	89	2	US-08-460-421A-13
23	56	50.9	89	2	US-08-717-169-7
24	56	50.9	89	4	US-09-228-901A-7
25	56	50.9	96	2	US-08-717-169-2
26	56	50.9	96	4	US-09-228-901A-2
27	56	50.9	109	1	US-07-829-462-1

Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 16, Appli
Sequence 1, Appli
Sequence 1, Appli
Patent No. 5455158
Patent No. 5455158
Sequence 12, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 26, Appli
Sequence 26, Appli
Sequence 26, Appli

28 56 50.9 109 1 US-08-340-812-1
29 56 50.9 109 1 US-08-459-064B-1
30 56 50.9 109 2 US-08-460-421A-1
31 56 50.9 109 2 US-08-717-169-1
32 56 50.9 109 4 US-09-228-901A-1
33 56 50.9 109 5 PCT-US93-00909-1
34 56 50.9 2231 1 US-08-153-799-16
35 56 50.9 2324 1 US-08-283-857-1
36 56 50.9 2324 5 PCT-US95-09819-1
37 56 50.9 2327 6 5455158-1
38 56 50.9 2327 6 5455158-1
39 56 50.9 2386 2 US-09-016-366A-12
40 56 50.9 2386 4 US-09-961-403-1
41 56 50.9 2446 2 US-08-551-356-2
42 56 50.9 2446 5 PCT-US93-12687-2
43 53 48.2 99 1 US-08-036-210-26
44 53 48.2 99 2 US-08-449-609-26
45 53 48.2 99 4 US-09-361-096A-26

ALIGNMENTS

RESULT 1
US-09-155-613A-2
; Sequence 2, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCES: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155.613A
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-155-613A-2
Query Match 50.9%; Score 56; DB 4; Length 26;
Best Local Similarity 100.0%; Pred.No. 0.087;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLWRP 10
| | | | | | | | | |
Db 15 ISKYLWRP 24

RESULT 2
US-09-155-613A-78
; Sequence 78, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCES: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155.613A
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005

```

; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotope
US-09-155-613A-78

Query Match          50.9%; Score 56; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.087;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ISKYILRWRP 10
Db      15 ISKYILRWRP 24

RESULT 3
US-07-829-462-2
; Sequence 2, Application US/07829462
; Patent No. 5453489
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki I.
; APPLICANT: Morla, Alex
; TITLE OF INVENTION: NOVEL FRAGMENTS OF FIBRONECTIN AND
; TITLE OF INVENTION: METHODS TO MODULATE THE ROLE OF FIBRONECTIN IN
; TITLE OF INVENTION: EXTRACELLULAR MATRIX ASSEMBLY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19920131
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,547
; REFERENCE/DOCKET NUMBER: P-LA 9179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-829-462-2

Query Match          50.9%; Score 56; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ISKYILRWRP 10
Db      8 ISKYILRWRP 17

RESULT 4
US-07-829-462-2
; Sequence 2, Application US/07829462
; Patent No. 5453489
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki I.
; APPLICANT: Morla, Alex
; TITLE OF INVENTION: NOVEL FRAGMENTS OF FIBRONECTIN AND
; TITLE OF INVENTION: METHODS TO MODULATE THE ROLE OF FIBRONECTIN IN
; TITLE OF INVENTION: EXTRACELLULAR MATRIX ASSEMBLY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19920131
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,547
; REFERENCE/DOCKET NUMBER: P-LA 9179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-829-462-2

Query Match          50.9%; Score 56; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ISKYILRWRP 10
Db      8 ISKYILRWRP 17

RESULT 5
US-08-459-064B-2
; Sequence 2, Application US/08459064B
; Patent No. 5747452
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: A METHOD OF MODULATING TUMOR CELL MIGRATION
; TITLE OF INVENTION: USING FIBRONECTIN TYPE III PEPTIDES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES LLP
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; 
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 7.39869 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKYLWRPVSPRNLYG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	20	2	AAI28902 MSP 1-1alp
2	110	100.0	642	8	ADQ39403 Human myo
3	110	100.0	642	8	ADR67316 Human bla
4	110	100.0	642	8	ADSL17489 Amino aci
5	110	100.0	642	8	ADR97658 Human fib
6	110	100.0	657	8	ADQ39409 Human myo
7	110	100.0	660	2	AAI28901 Human mig
8	61	55.5	2265	8	ADSL17498 Amino aci
9	61	55.5	2265	8	ADR97667 Bovine fi
10	56	50.9	20	2	AAI28911 Peptide e
11	56	50.9	26	2	AAW75324 Human fib
12	56	50.9	26	2	AAW75260 Human fib
13	56	50.9	26	2	AAW70995 Motif 1 o
14	56	50.9	31	2	AAW19970 Fibronect
15	56	50.9	31	2	AAW57192 Fibronect
16	56	50.9	31	2	AAW82986 Human fib
17	56	50.9	56	2	AAW57194 Fibronect
18	56	50.9	66	2	AAW19973 Fibronect
19	56	50.9	66	2	AAW82984 Human fib
20	56	50.9	75	2	AAW57191 Fibronect
21	56	50.9	75	6	ABU63426 Human fib
22	56	50.9	75	8	ADI58820 Angiogene
23	56	50.9	85	2	AAW19972 Fibronect
24	56	50.9	85	2	AAW82983 Human fib
25	56	50.9	89	2	AAW19974 Fibronect

26	56	50.9	89	2	AAW57195	AAW57195 Fibronect
27	56	50.9	89	2	AAW82985	AAW82985 Human fib
28	56	50.9	96	2	AAW57190	AAW57190 Fibronect
29	56	50.9	109	2	AAW39729	AAW39729 Fiset typ
30	56	50.9	109	2	AAW57189	AAW57189 Fibronect
31	56	50.9	109	2	AAW82982	AAW82982 Human fib
32	56	50.9	720	2	AAI28914	AAI28914 Fibronect
33	56	50.9	984	8	ADQ39406	ADQ39406 Human myo
34	56	50.9	993	2	AAW80839	AAW80839 Cathepsin
35	56	50.9	1173	4	ABG22275	ABG22275 Novel hum
36	56	50.9	1179	8	ADP75952	ADP75952 Human min
37	56	50.9	1189	2	AAW80839	AAW80839 Human ser
38	56	50.9	1223	8	ADP30365	ADP30365 Human sec
39	56	50.9	1359	8	ADP75957	ADP75957 Human leu
40	56	50.9	1381	2	AAW80832	AAW80832 Human ser
41	56	50.9	2182	8	ADR66462	ADR66462 Human pro
42	56	50.9	2182	8	ADR66120	ADR66120 Human pro
43	56	50.9	2220	6	ABO01289	ABO01289 Human pro
44	56	50.9	2220	8	ADN95950	ADN95950 Human NOV
45	56	50.9	2265	4	AAW38647	AAW38647 Human pol

ALIGNMENTS

RESULT 1

AAI28902
ID AAI28902 standard; peptide; 20 AA.

XX AAI28902;

XX 21-SEP-1999 (first entry)

XX MSF 1-alpha peptide epitope.

XX Migration stimulatory factor; MSF; cell migration; modulation; human;

XX wound healing; scarring; MSF1-alpha; epitope; fibronectin.

XX Synthetic.

XX Homo sapiens.

XX W09931233-A1.

XX 24-JUN-1999.

XX 15-DEC-1998; 98WO-GB003766.

XX 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI; 1999-430039/36.

Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

Claim 18; Page 57; 86pp; English.

The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. Sequences AAI28902-906 represent peptide epitopes of MSF against which monoclonal antibodies that are specific to MSF and do not cross-react with fibronectin are raised

SQ Sequence 20 AA;

Query Match 100.0%; Score 110; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.3e-09;

XX	SQ	Sequence 642 AA;
	Query Match	100.0%; Score 110; DB 8; Length 642;
	Best Local Similarity	100.0%; Pred. No. 5.3e-08;
	Matches 20; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ISKYILWRPVSIPPNLGY 20 	
Db	623 ISKYILWRPVSIPPNLGY 642	
RESULT 3		
ADR67316	ID	ADR67316 standard; protein; 642 AA.
XX	AC	ADR67316;
XX	DT	02-DEC-2004 (first entry)
XX	DE	Human bladder cancer associated amino acid sequence.
XX	KW	bladder cancer tissue; bladder cancer; cytostatic.
XX	OS	Homo sapiens.
XX	PN	WO2004076613-A2.
XX	PD	10-SEP-2004.
XX	PF	24-FEB-2004; 2004WO-D8000364.
XX	PR	26-FEB-2003; 2003DB-01009729.
XX	PA	(HERR/) HERR A. (HINZ/) HINZMANN B. (DAHL/) DAHL E. (STAU/) STAU E. (PILA/) PILARSKY C. (SPEC/) SPECHT T.
PI	Herr A, Hinzmann B, Dahl E, Staub E, Pilarsky C, Specht T;	
XX	WPI; 2004-653385/63.	
XX	New nucleic acids, and encoded proteins, from bladder cancer tissue, useful for diagnosis, treatment and in screening for specific binding agents.	
XX	Claim 2; Fig 3; 112pp; German.	
CC	The present invention describes nucleic acids (I) associated with bladder cancer tissue. Also described: (1) peptides and proteins (II) containing an amino acid sequence encoded by (I); (2) a method for diagnosing bladder cancer (BC), or monitoring its progression, that uses (I), (II) or agents (Z) that inhibit, or bind to, (I) or (II), labelled with a reporter; and (3) a method for treating BC that uses (I), (II) or (Z). (I) and (II) have cytostatic activity. (I) and (II) can be used to detect (and monitor progression of) bladder cancer (BC), or the risk of developing it; to screen for specific binding agents (Z), and to treat BC. (Z) are also useful as diagnostic and therapeutic agents. The present sequence represents a human amino acid sequence associated with bladder cancer, which is used in the exemplification of the present invention.	
XX	Sequence 642 AA;	
	Query Match	100.0%; Score 110; DB 8; Length 642;
	Best Local Similarity	100.0%; Pred. No. 5.3e-08;
	Matches 20; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ISKYILWRPVSIPPNLGY 20 	
Db	623 ISKYILWRPVSIPPNLGY 642	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 415.948 Seconds
(without alignments)
1830.243 Million cell updates/sec

Title: US-09-581-651D-5
Perfect score: 110
Sequence: 1 ISKYLWRPVSIPPNLGY 20

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2_1/USPTO.spool_b/US09581651/runat_07112005_092223_28789/app_query.fasta_1.1834
-DB=EST -QFMT=fastap -SURFIX=p2n.rst -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -THR ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651.cgm 1 1 8180 @runat_07112005_092223_28789 -NCPU=3 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hcc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g861:*
9: gb_g862:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	110	100.0	362	1	AI263888 qi08d11.x
C 2	110	100.0	391	1	AA234360 zr72c02.s
C 3	110	100.0	440	1	AI754331 cr23e08.x
C 4	110	100.0	670	5	BU620723 UI-H-FL1-
C 5	110	100.0	706	4	BM674187 UI-E-EUO-
C 6	110	100.0	747	6	CA423317 UI-H-PF1-
C 7	110	100.0	967	5	EX402381 BX402381
C 8	110	100.0	1044	5	BX398837 BX398837
C 9	110	100.0	1052	5	BX380582 BX380582

10	65	59.1	658	2	BB625874	BB625874
11	61	55.5	532	5	BQ830132	BQ830132
12	61	55.5	546	5	BQ830316	LL6in2173
13	61	55.5	562	2	BB709173	BB709173
14	61	55.5	702	9	CE165323	CE165323
15	60	54.5	348	2	BB812986	BB812986
c 16	60	54.5	397	1	AJ665984	AJ665984
c 17	60	54.5	607	7	CK450361	CK450361
18	60	54.5	613	9	CL586124	CL586124
19	60	54.5	755	7	CK451343	CK451343
20	60	54.5	841	4	BI684770	BI684770
c 21	60	54.5	886	7	CNI63427	CNI63427
22	60	54.5	897	7	CNI62046	CNI62046
23	59	53.6	192	2	BB361047	BB361047
24	59	53.6	1128	8	CC193783	CC193783
25	58	52.7	226	2	BB394322	BB394322
c 26	58	52.7	1133	4	BM478486	BM478486
c 27	58	52.7	1521	7	CK411745	CK411745
28	57.5	52.3	423	2	BB812942	BB812942
29	57.5	52.3	828	1	AI287222	AI287222
c 30	57	51.8	507	8	BZ695434	BZ695434
c 31	57	51.8	890	2	BE789442	BE789442
c 32	56.5	51.4	707	1	AU140730	AU140730
c 33	56	50.9	210	4	BI040656	BI040656
c 34	56	50.9	219	4	BI054371	BI054371
c 35	56	50.9	220	4	BI040247	BI040247
c 36	56	50.9	260	2	BF880844	BF880844
c 37	56	50.9	268	2	BF367054	BF367054
c 38	56	50.9	269	2	BF369623	BF369623
c 39	56	50.9	295	2	BE812941	BE812941
40	56	50.9	301	1	AA853208	AA853208
41	56	50.9	316	2	BF876014	BF876014
c 42	56	50.9	320	1	AA302280	AA302280
c 43	56	50.9	320	5	BQ319620	BQ319620
44	56	50.9	328	2	BF869573	BF869573
c 45	56	50.9	334	2	BF808854	BF808854

ALIGNMENTS

RESULT 1
AI263888/c

LOCUS
DEFINITION

AI263888 qi08d11.x1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:1855893
3', similar to gb:X02761_cdsl FIBRONECTIN PRECURSOR (HUMAN);, mRNA
sequence.

ACCESSION
VERSION

AI263888
AI263888.1 GI:3872091

KEYWORDS
SOURCE

Homo sapiens
Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.

FEATURES

source

1..362

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="IMAGE:1855893"

/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"

/lab_hosts="DH10B"

/clone_lib="Soares_NHMPu_S1"

/note="Organ: mixed (see below); Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NbHPU, and fetal heart NbHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

ORIGIN

Alignment Scores: 2.23e-06 Length: 362
 Pred. No.: 110.00 Matches: 20
 Score: 110.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-09-581-651D-5 (1-20) x A1263888 (1-362)

Qy 1 IleSerLysTyrIleLeuArgTrrpArgProValSerIleProProArgAsnLeuGlyTyr 20
 |||||
 Db 214 ATTTCAGTACATCTCAGGTGGAGACTGTGAGTATCCACCCAGAAACCTTGGATAC 155

RESULT 2

AA234360/c

LOCUS

DEFINITION AA234360 391 bp mRNA linear EST 06-AUG-1997
 zr72c02.sl Soares NHMPu S1 Homo sapiens cDNA clone IMAGE:668930 3',
 similar to gb:X02761_cds1 FIBROBLAST PRECURSOR (HUMAN);, mRNA
 sequence.

ACCESSION

AA234360

VERSION

AA234360.1 GI:1858635

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
 White, Y., Wyllie, T., Waterson, K. and Wilson, R.

WashU-Merck EST Project 1997

Unpublished (1997)

CONTACT: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 678 Std Error: 0.00

Seq primer: -41m3 fwd. ET from Amersham

High quality sequence stop: 376.

Location/Qualifiers

1..391

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:5562897"

/db_xref="taxon:9606"

/clone="IMAGE:668930"

/tissue_type="Pooled human melanocyte, fetal heart, and

pregnant uterus"

/lab_host="DH108"

/clone_lib="Soares NHMPu S1"

/note="Organ: mixed (see below); Vector: pTT3D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; Equal amounts of plasmid DNA from three

normalized libraries (melanocyte 2NbHM, pregnant uterus

NbHPU, and fetal heart NbHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

ORIGIN

Alignment Scores: 2.44e-06 Length: 391
 Pred. NO.: 110.00 Matches: 20
 Score: 110.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-09-581-651D-5 (1-20) x AA234360 (1-391)

Qy 1 IleSerLysTyrIleLeuArgTrrpArgProValSerIleProProArgAsnLeuGlyTyr 20
 |||||
 Db 205 ATTTCAGTACATCTCAGGTGGAGACTGTGAGTATCCACCCAGAAACCTTGGATAC 146

RESULT 3

AI754331/c

LOCUS

DEFINITION AI754331 440 bp mRNA linear EST 20-JUN-2002
 cr23e08.xl Human bone marrow stromal cells Homo sapiens cDNA clone
 HMSC cr23e08 3', mRNA sequence.

ACCESSION

AI754331

VERSION

AI754331.1 GI:5132595

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

Jia, L., Young, M.F., Powell, J., Yang, L., Ho, N.C., Hotchkiss, R.,
 Robey, P.G. and Francomano, C.A.

Gene expression profile of human bone marrow stromal cells:
 high-throughput expressed sequence tag sequencing analysis

Genomics 79 (1), 7-17 (2002)

21686149

11827452

PUBMED

COMMENT

Contact: Libin Jia

Medical Genetics Branch

National Human Genome Research Institute

10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA

Tel: 301-402-4877

Fax: 301-496-7157

Email: libin@helix.nih.gov

DNA Sequencing and analyses by National Institutes of Health

Intramural Sequencing Center (NISC).

Plate: 23 row: e column: 08

Seq primer: -21M13 forward primer (ABI).

Location/Qualifiers

1..440

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="HMSC_cr23e08"

/sex="mixed"

/tissue_type="bone marrow stroma"

/dev_stage="mixed"

/lab_host="XLI-Blue MRF/SOLR"

/clone_lib="Human bone marrow stromal cells"

/note="Vector: pBluescript; Site 1: EcoRI; Site 2: XhoI;

mRNA made from human bone marrow stroma, cDNA made by

oligo-dT priming. Directionally cloned. Size-selected for

average insert size >0.5 kb. Library constructed by Dr.

Marian Young and Dr. Pamela Gehron Robey (NIDCR). Library

supplied by Dr. Libin Jia (NHGRI)"

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 102.693 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-5
Perfect score: 110
Sequence: 1 ISKYLWRVSPRPRLNGY 20

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 1989580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool.h/US09581651/runat_07112005.092226.28867/app_query.fasta.1.1834
-DB=Published Applications NA -QFMT=fastp -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blseqm62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09581651 @CGN 1 1 1041 @runat_07112005.092226.28867
-NCPU=6 -ICPU=1 -NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq*
23: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	471	18	US-10-242-535A-42467
2	110	100.0	471	19	US-10-085-783A-42467
3	110	100.0	600	22	US-10-956-157-9523
4	110	100.0	2127	17	US-10-210-120-49
5	110	100.0	2127	22	US-10-956-157-4288
6	110	100.0	2127	24	US-10-909-035-49
7	110	100.0	2443	20	US-10-741-601-70
8	110	100.0	2443	22	US-10-741-600-238
9	110	100.0	2488	20	US-10-741-601-75
10	110	100.0	2488	22	US-10-741-600-244
11	110	100.0	4295	18	US-10-144-194A-51
12	110	100.0	4295	21	US-10-491-566-51
13	110	100.0	87467	20	US-10-741-601-5634
14	110	100.0	87467	22	US-10-741-600-17824
15	59	53.6	937	18	US-10-369-493-27562
c 16	56	50.9	540	22	US-10-695-140-1
17	56	50.9	567	20	US-10-767-701-30779
c 18	56	50.9	1892	20	US-10-384-339C-133
c 19	56	50.9	2055	24	US-10-974-558-1
c 20	56	50.9	3522	24	US-10-450-763-22266
c 21	56	50.9	6189	16	US-10-259-273-41
c 22	56	50.9	6189	26	US-11-006-313-41
23	56	50.9	6510	20	US-10-741-601-72
24	56	50.9	6510	22	US-10-741-600-241
25	56	50.9	6988	19	US-10-236-392-1
26	56	50.9	7361	19	US-10-236-392-3
27	56	50.9	7677	22	US-10-956-157-4995
28	56	50.9	7679	22	US-10-831-704-38
29	56	50.9	7680	9	US-09-964-824A-574
30	56	50.9	7680	15	US-10-171-311-63
31	56	50.9	7680	18	US-10-236-031B-69
32	56	50.9	7680	18	US-10-374-979-75
33	56	50.9	7680	19	US-10-182-936A-75
34	56	50.9	7680	19	US-10-641-643-1289
35	56	50.9	7680	20	US-10-717-597-222
36	56	50.9	7680	20	US-10-788-792-79
37	56	50.9	7680	21	US-10-477-238A-654
38	56	50.9	7680	21	US-10-580-287A-654
39	56	50.9	7680	22	US-10-278-698-88
40	56	50.9	7680	22	US-10-278-698-603
41	56	50.9	7680	22	US-10-843-641A-5877
42	56	50.9	7680	22	US-10-477-173-654
43	56	50.9	7680	22	US-10-852-135A-52
44	56	50.9	7705	18	US-10-447-161-4
45	56	50.9	7795	16	US-10-084-817-2

ALIGNMENTS

RESULT 1
US-10-242-535A-42467
; Sequence 42467, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Lew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242.535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42467
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-42467

Alignment Scores:
Pred. No.: 5,23e-08 Length: 471
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 18

US-09-581-651D-5 (1-20) x US-10-242-535A-42467 (1-471)

Qy 1 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 363 ATTCCAAAGTACATTCTCAGGTGGAGACCTGTGAGTATCCACCAGAAACCTTGGATAC 422

RESULT 2

US-10-085-783A-42467
; Sequence 42467, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis

; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42467
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-42467

Alignment Scores:
Pred. No.: 5,23e-08 Length: 471
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 19

US-09-581-651D-5 (1-20) x US-10-085-783A-42467 (1-471)

Qy 1 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 363 ATTCCAAAGTACATTCTCAGGTGGAGACCTGTGAGTATCCACCAGAAACCTTGGATAC 422

RESULT 3

US-10-956-157-9523
; Sequence 9523, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William

; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9523
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-9523

Alignment Scores:
Pred. No.: 6.8e-08 Length: 600
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 22

US-09-581-651D-5 (1-20) x US-10-956-157-9523 (1-600)

Qy 1 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 396 ATTCCAAAGTACATTCTCAGGTGGAGACCTGTGAGTATCCACCAGAAACCTTGGATAC 455

RESULT 4

US-10-210-120-49
; Sequence 49, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:
; APPLICANT: Chinnaiyan, Arul M.

; APPLICANT: Rubin, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UM-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/334,468
; PRIOR FILING DATE: 2001-11-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-210-120-49

Alignment Scores:
Pred. No.: 2.69e-07 Length: 2127
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 17

US-09-581-651D-5 (1-20) x US-10-210-120-49 (1-2127)

Qy 1 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 1923 ATTCCAAAGTACATTCTCAGGTGGAGACCTGTGAGTATCCACCAGAAACCTTGGATAC 1982

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:49:16 ; Search time 20.1307 Seconds
(without alignments)
1625.652 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKIILWRVSIIPRNGLY 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/h/US09581651/runat_07112005_092224_28801/app.query.fasta_1.1834
-DB=issued_Patents_NA -QPMR=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651@cgn 1.187 @runat_07112005_092224_28801 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	57	51.8	374159	4	US-09-949-016-15868
2	56	50.9	186	1	Sequence 15868, A
3	56	50.9	7679	4	US-08-153-799-5
4	56	50.9	7680	4	Sequence 5, Appli
5	56	50.9	7680	5	Sequence 38, Appl
6	56	50.9	7705	2	Sequence 1289, Ap
7	56	50.9	7705	1	Sequence 6, Appli
8	56	50.9	7705	6	Sequence 16, Appl
9	56	50.9	7705	6	Sequence 16, Appl
10	56	50.9	7803	2	US-08-551-356-1
11	56	50.9	7803	5	Patent No. 5455158
C 12	56	50.9	7989	4	Sequence 1, Appli
					Sequence 10, Appl

C 13	56	50.9	8001	4	US-09-539-601-7	Sequence 7, Appli
C 14	56	50.9	8001	4	US-09-539-601-16	Sequence 16, Appl
C 15	56	50.9	8001	4	US-09-539-601-22	Sequence 22, Appl
C 16	56	50.9	8001	4	US-09-539-601-28	Sequence 28, Appl
C 17	56	50.9	8044	4	US-09-566-921-135	Sequence 135, App
C 18	56	50.9	8637	4	US-09-539-601-4	Sequence 4, Appli
C 19	56	50.9	8638	4	US-10-029-907-6	Sequence 6, Appli
C 20	56	50.9	8638	4	US-10-029-907-7	Sequence 7, Appli
C 21	56	50.9	8638	4	US-10-029-907-24	Sequence 24, Appl
C 22	56	50.9	8638	4	US-10-029-907-25	Sequence 25, Appl
C 23	56	50.9	8639	4	US-10-029-907-1	Sequence 1, Appli
C 24	56	50.9	8642	4	US-10-029-907-2	Sequence 2, Appli
C 25	56	50.9	8643	4	US-10-029-907-4	Sequence 4, Appli
C 26	56	50.9	8648	4	US-10-029-907-5	Sequence 5, Appli
C 27	56	50.9	8649	4	US-09-539-601-13	Sequence 13, Appl
C 28	56	50.9	11076	4	US-09-539-601-1	Sequence 1, Appli
C 29	56	50.9	11076	4	US-09-539-601-19	Sequence 19, Appl
C 30	56	50.9	11076	4	US-09-539-601-25	Sequence 25, Appl
C 31	56	50.9	11076	4	US-09-539-601-31	Sequence 31, Appl
C 32	54	49.1	7475	2	US-08-971-036-1	Sequence 1, Appli
C 33	54	49.1	7475	3	US-09-096-570-1	Sequence 1, Appli
C 34	54	49.1	7475	3	US-09-265-617B-1	Sequence 1011, Ap
C 35	53.5	48.6	5137	4	US-09-949-016-1011	Sequence 3107, Ap
C 36	53.5	48.6	5140	4	US-09-949-016-3107	Sequence 12753, A
C 37	53.5	48.6	9681	4	US-09-949-016-12753	Sequence 14849, A
C 38	53.5	48.6	9681	4	US-09-949-001-203	Sequence 203, App
C 39	53	48.2	601	4	US-09-949-001-477	Sequence 477, App
C 40	53	48.2	601	4	US-09-949-001-477	Sequence 10, Appl
C 41	53	48.2	2320	3	US-09-202-904A-10	Sequence 1, Appli
C 42	53	48.2	4975	2	US-08-249-687C-1	Sequence 1, Appli
C 43	53	48.2	4989	2	US-08-566-392A-3	Sequence 3, Appli
C 44	53	48.2	4989	2	US-08-625-819-1	Sequence 1, Appli
C 45	53	48.2	4989	3	US-08-755-558-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-949-016-15868/c
; Sequence 15868, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15868
; LENGTH: 374159
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15868

Alignment Scores:	1.46e+04	Length:	374159
Pred. No.:	57.00	Matches:	11
Score:	73.33%	Conservative:	0
Percent Similarity:	73.33%	Mismatches:	4
Best Local Similarity:	73.33%	Indels:	0
Query Match:	51.82%	Gaps:	0
DB:	4		
US-09-581-651D-5 (1-20) x US-09-949-016-15868 (1-374159)			
Qy	2 SerLysTyrIleLeuArgTrpArgProValSerIleProProArg 16		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:30:56 ; Search time 65.5686 Seconds
(without alignments)
1805.663 Million cell updates/sec

Title: US-09-581-651D-5
Perfect score: 110
Sequence: 1 ISKYLWRPVSIPPRNLGY 20

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092222_28772/app_query_fasta_1.1834
-DB=N_Geneseq -QFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651.qcgn 1.1 1052 @runat_07112005_092222_28772 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOBURY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database :

N_Geneseq 16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	1929	13 ADS17488	Adsl7488 Nucleotid
2	110	100.0	1929	13 ADS97657	Adr97657 Human fib
3	110	100.0	2127	10 ADD18477	Add18477 Human pro
4	110	100.0	2147	2 AAX81299	Aax81299 Human mig
5	110	100.0	2358	13 ADR67201	Adr67201 Human bla

RESULT 1
ADS17488
ID ADS17488 standard; DNA; 1929 BP.
XX
AC ADS17488;
XX
DT 02-DEC-2004 (first entry)
XX
DE Nucleotide sequence of human fibronectin 1.
XX
KW cell state; time-lapse profile; protein-protein interaction;
KW drug screening; cancer; infectious disease; allergy; hypertension;
KW hyperlipaemia; diabetes; cardiac disease; cerebral infarction; dementia;
KW obesity; arteriosclerosis; infertility; mental disease; nervous disease;
KW cataract; progeria; hypersensitivity; ultraviolet radiation; human;
KW fibronectin 1; actin acting substance; transfection array; gene; db.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1929
FT /tag= a
FT /product= "fibronectin 1"
XX
FN WO2004079007-A2.
XX
PD 16-SEP-2004.
XX
PF 03-MAR-2004; 2004WO-JP002694.

ALIGNMENTS

6	110	100.0	2443	13	ADQ38575	Adq38575 Human SNP
7	110	100.0	2488	13	ADQ38581	Adq38581 Human SNP
8	110	100.0	4295	8	ACC72037	Acc72037 BCU0770 g
9	110	100.0	11026	4	AAK72958	Aak72958 Human lmm
10	110	100.0	31749	4	AAK72959	Aak72959 Human lmm
11	110	100.0	78925	3	AAC89888	Aac89888 Human FN
12	59	53.6	937	13	ADS49132	Ads49132 Bacterial
13	56	50.9	225	10	ADD93721	Add93721 Hepatitis
14	56	50.9	499	10	ACD97182	Acd97182 Human col
15	56	50.9	540	12	ADN96854	Adn96854 Hepatitis
16	56	50.9	1892	6	ABV78249	Abv78249 Hepatitis
17	56	50.9	1892	6	ABZ35825	Abz35825 Hepatitis
18	56	50.9	1892	6	ABX10068	Abx10068 HCV NS3 D
19	56	50.9	1892	6	ABL91790	AbL91790 HCV polyn
20	56	50.9	3522	5	AAS86462	Aas86462 DNA encod
21	56	50.9	3540	12	ADP75954	Adp75954 Human min
22	56	50.9	3669	12	ADP29134	Adp29134 Human sec
23	56	50.9	4080	12	ADP75958	Adp75958 Human leu
24	56	50.9	4860	3	AAA35009	Aaa35009 Human ade
25	56	50.9	6189	10	ADD67965	Add67965 Hepatitis
26	56	50.9	6510	13	ADQ38578	Adq38578 Human SNP
27	56	50.9	6816	10	ABX74443	Abx74443 Human CDN
28	56	50.9	6988	9	ACD06169	Acd06169 Human CDN
29	56	50.9	7020	13	ADR90518	Adr90518 Human CDN
30	56	50.9	7049	12	ADN95947	Adn95947 Human NOV
31	56	50.9	7141	6	AAD25333	Aad25333 Hepatitis
32	56	50.9	7242	13	ADR66637	Adr66637 Human pro
33	56	50.9	7242	13	ADR65953	Adr65953 Human pro
34	56	50.9	7361	9	ACD06170	Acd06170 Human CDN
35	56	50.9	7361	12	ADN95949	Adn95949 Human NOV
36	56	50.9	7550	8	ACC00412	Acc00412 Human cell
37	56	50.9	7677	13	ACN37820	Acn37820 Tumour-as
38	56	50.9	7679	10	ADB31322	Adb31322 Testoster
39	56	50.9	7680	2	AAT17551	Aat17551 Human fib
40	56	50.9	7680	3	AAF21131	Aaf21131 Human low
41	56	50.9	7680	5	ABA82689	Aba82689 Fibronect
42	56	50.9	7680	6	ABL67540	AbL67540 Thyroid c
43	56	50.9	7680	6	ABT11082	Abt11082 Human bre
44	56	50.9	7680	8	ACF03878	Acf03878 Human fib
45	56	50.9	7680	8	ABX10391	Abx10391 DNA encod

```
XX PR 04-MAR-2003; 2003JP-00057870.
XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX PI Miyake M, Yoshikawa T, Uchimura E, Miyake J;
XX DR WPI; 2004-662438/64.
XX DR P-PSDB; ADS17489.
XX PT Presenting a state of a cell, useful for diagnosing and treating a
XX PT disease, e.g. cancer, infectious disease, allergy, diabetes, dementia,
XX PT obesity, infertility, or cataract, comprises obtaining a time-lapse
XX PT profile of the cell.
XX PS Disclosure; SEQ ID NO 1; 532pp; English.
XX CC The specification describes a method and system for accurately presenting
XX CC a state of a cell. The method comprises obtaining a time-lapse profile of
XX CC the cell by time-lapse monitoring of a gene state associated with at
XX CC least one gene derived from the cell, and presenting the time-lapse
XX CC gene state includes expression of the gene. The method and system are
XX CC useful presenting a state of a cell. The method can allow the elucidation
XX CC of key protein-protein interactions suitable for targeting by drug
XX CC screening protocols. The method is useful for diagnosing or treating a
XX CC disease, e.g. cancer, infectious disease due to viruses or bacteria,
XX CC allergy, hypertension, hyperlipaemia, diabetes, cardiac disease, cerebral
XX CC infarction, dementia, obesity, arteriosclerosis, infertility, mental and
XX CC nervous diseases, cataract, progeria, or hypersensitivity to ultraviolet
XX CC radiation. The present sequence encodes human fibronectin 1. Bovine
XX CC fibronectin was used as a candidate for an actin acting substance. The
XX CC actin acting substance was used with transfection reagents and amplified
XX CC plasmid DNA in assays using transfection arrays, in the course of the
XX CC invention.
XX SQ Sequence 1929 BP; 507 A; 462 C; 544 G; 416 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.62e-07 Length: 1929
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-581-651D-5 (1-20) x ADS17488 (1-1929)
Qy 1 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 1867 ATTTCGAAGTACATTCCTCAGGTGGAGACCTGTGAGTATCCCAACCGAATCCTTGGATAC 1926

RESULT 2
ADR97657
ID ADR97657 standard; DNA; 1929 BP.
XX AC ADR97657;
XX DT 02-DEC-2004 (first entry)
XX DE Human fibronectin 1 DNA, an actin acting substance SeqID 1.
XX KW human; gene; ds; transfection efficiency; actin acting substance;
XX KW extracellular matrix; fibronectin 1; gene introduction reagent.
XX OS Homo sapiens.
XX FH Location/Qualifiers
XX FT 1..1929
XX FT /*tag= a
XX FT /product= "Fibronectin protein"
XX PN WO2004079332-A2.

16-SEP-2004.
03-MAR-2004; 2004WO-JP002696.
04-MAR-2003; 2003JP-00057869.
(NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
Miyake M, Yoshikawa T, Uchimura E, Miyake J;
WPI; 2004-677173/66.
P-PSDB; ADR97658.
Composition comprising an actin acting substance or an actin acting
substance and a target substance, useful for increasing the efficiency of
introducing a target substance into a cell and in cell biology or genetic
engineering.
Claim 4; SEQ ID NO 1; 347pp; English.
This invention relates to a novel composition and method for increasing
the efficiency of introducing a target substance into a cell.
Specifically, it refers to the introduction of DNA (e.g. the gene of
interest to be transfected), polypeptides, sugars or complexes thereof
into a cell, and comprises an actin acting substance. The present
invention describes the actin acting substance as an extracellular matrix
protein, a variant or fragment thereof selected from fibronectin, laminin
or vitronectin. The composition further comprises a gene introduction
reagent selected from cationic polymers, cationic lipids, and calcium
phosphate, as well as a gold colloid particle that is contacted with the
cell. As such, the composition, kit, device or method is useful for
increasing the efficiency of introducing a target substance into a cell
and thus is useful in the fields of cell biology, genetic engineering and
molecular biology. This polynucleotide sequence is the human fibronectin
DNA sequence of the invention.
SQ Sequence 1929 BP; 507 A; 462 C; 544 G; 416 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.62e-07 Length: 1929
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-581-651D-5 (1-20) x ADR97657 (1-1929)
Qy 1 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 1867 ATTTCGAAGTACATTCCTCAGGTGGAGACCTGTGAGTATCCCAACCGAATCCTTGGATAC 1926

RESULT 3
ADD18477
ID ADD18477 standard; DNA; 2127 BP.
XX AC ADD18477;
XX DT 15-JAN-2004 (first entry)
XX DE Human prostate cancer diagnosis related DNA sequence SeqID49.
XX KW prostate tissue; cancer diagnostic; cancer marker; prostate cancer; PCA;
XX KW male cancer-related death; serum biomarker; tissue biomarker; cytostatic;
XX KW gene therapy; prostate biopsy tissue; ANAC;
XX KW alpha-methylacyl-coenzyme A racemase; diagnosing cancer; cell growth;
XX KW human; ds.
XX OS Homo sapiens.
XX PN WO2003012067-A2.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 535.477 Seconds
(without alignments)
1809.796 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKYLRWPRVSPRNLGY 20

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool_h/US09581651/runat_07112005_092223_28778/app_query.fasta_1.1834
-DB=GenEmbl -QPM=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CGN 1.1 9936 @runat_07112005_092223_28778 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	110	100.0	471	6	C0697541 Sequence
2	110	100.0	1929	6	C0871810 Sequence
3	110	100.0	1929	6	C0871828 Sequence
4	110	100.0	2147	6	BD137021 Polypepti

5	110	100.0	2147	6	AX003229	Sequence
6	110	100.0	2147	9	HS276395	Homo sapi
7	110	100.0	2192	9	HS276395	Homo sapi
8	110	100.0	2358	6	Q0875358	Sequence
9	110	100.0	4321	9	HS276395	Homo sapi
10	110	100.0	143947	9	AC012462	Homo sapi
11	65	59.1	197121	2	AC127101	Rattus no
12	65	59.1	229287	2	AC107509	Rattus no
13	65	59.1	229287	2	AC107509	Rattus no
14	65	59.1	244185	2	AC123496	Rattus no
15	65	59.1	255929	2	AC095457	Rattus no
16	62	56.4	7085	6	A76577	Sequence 9
17	62	56.4	231764	2	AC119470	Rattus no
18	62	56.4	233445	2	AC129682	Rattus no
19	61	55.5	185651	2	AC124821	Mus muscu
20	61	55.5	204737	2	AC091456	Mus muscu
21	60	54.5	163319	2	AC110512	Mus muscu
22	60	54.5	198224	2	AC141795	Apis mell
23	59.5	54.1	287860	1	AE017265	Bacillus
24	59	53.6	543	14	AF369258	Hepatitis
25	58.5	53.2	180594	10	AC112270	Mus muscu
26	58.5	53.2	267977	2	AC151721	Mus muscu
27	58	52.7	9427	14	HPCJRNA	Hepatitis C
28	58	52.7	11235	1	AE008320	Agrobacte
29	58	52.7	11449	1	AE015266	Shigella
30	58	52.7	14357	1	AE009292	Agrobacte
31	58	52.7	34670	3	LMFL1648	Leishmani
32	58	52.7	110000	2	LMFLCHR32_09	Continuatio
33	58	52.7	140566	2	AC141532	Rattus no
34	58	52.7	215689	2	AC111687	Rattus no
35	58	52.7	292088	1	AE016986	Shigella
36	57	51.8	210	14	HPCNS3A2	Hepatitis C
37	57	51.8	3640	14	AY070174	Hepatitis
38	57	51.8	9365	14	AY587844	Hepatitis
39	57	51.8	78184	9	AC067722	Homo sapi
40	57	51.8	80938	2	AC023502	Homo sapi
41	57	51.8	106556	9	HSDJ512E2	Human DNA
42	57	51.8	109853	8	AC135417	Oryza sat
43	57	51.8	141415	2	AC008032	Homo sapi
44	57	51.8	144014	2	AC149702	Bos tauru
45	57	51.8	175916	2	AC007747	Homo sapi

ALIGNMENTS

RESULT 1
C0697541
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

C0697541
Sequence 42467 from Patent WO02070737.
C0697541
C0697541.1 GI:42248968

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
Liew C.C., Marshall, W.B. and Zhang, H.

Compositions and methods relating to osteoarthritis

Patent: WO 02070737-A 42467 12-SEP-2002;

Chondrogene Inc. (CA)

Location/Qualifiers

1..471

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred No.: 5e-08

Score: 110.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Length: 471

Matches: 20

Conservative: 0

Mismatches: 0

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2005, 22:29:25 ; Search time 7025 Seconds
(without alignments)
11633.308 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 2147

Sequence: 1 caaacttggtggaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 segs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1885.8	87.8	7777	3	CR749316 Homo sapi
2	1885.8	87.8	7885	3	CR749281 Homo sapi
3	1882.6	87.7	8121	3	CR749317 Homo sapi
4	1875.4	87.3	7501	3	BC078656 Homo sapi
5	1460.2	68.0	8315	3	AK090135 Mus muscu
6	1391.2	64.8	8329	3	AK090130 Mus muscu
7	908.6	42.3	1044	5	EX398837 BX398837
8	875.2	40.8	1052	5	EX380582 BX380582
9	794.8	37.0	943	5	EX391752 BX391752
10	784	36.5	1026	5	EX386270 BX386270
11	766.6	35.7	816	1	AU141008 AU141008
12	766.2	35.7	967	5	EX402381 BX402381
13	749	34.9	766	7	CN419594 170004245
14	748.8	34.9	1050	5	EX417945 BX417945
15	733.2	34.1	765	4	BM715855 UI-E-EJ0-
16	731	34.0	861	1	AU140971 AU140971
17	727.4	33.9	737	1	AU140993 AU140993
18	723.2	33.7	731	1	AL140526 AU140526
19	718.2	33.5	739	1	AL1706215 DKFZp686J
20	707.2	32.9	747	6	CA423317 UI-H-PP1-
21	700.2	32.6	730	1	AU140834 AU140834
22	693	32.3	695	7	CN419477 CN419477
23	693	32.3	706	4	BM674187 UI-E-EJ0-
24	686.8	32.0	2107	3	AK054456 Mus muscu

25 686.6 32.0 837 1 AU140973
26 680.8 31.7 699 1 AU140889
27 676.6 31.5 734 1 AU140802
28 676 31.5 705 1 AL603362
29 673.2 31.4 705 1 AU140789
30 669 31.2 723 7 CN419622
31 664 30.9 667 1 AU140910
32 663.2 30.9 911 5 BX398838
33 662 30.8 738 6 CD613781
34 660.6 30.8 670 5 BU620723
35 656.4 30.6 835 7 CN538822
36 655.4 30.5 1038 5 BX439175
37 655.2 30.5 673 1 AU140450
38 654 30.5 851 1 AU141024
39 652 30.4 859 7 CF616056
40 646.4 30.1 796 1 AU140814
41 639.8 29.8 699 1 AL706221
42 635 29.6 748 1 AU140948
43 632.2 29.4 748 1 AU140991
44 628.6 29.3 758 1 AU140556
45 626.4 29.2 673 7 CN419479

ALIGNMENTS

RESULT 1
CR749316
LOCUS CR749316 7777 bp mRNA linear HTC 19-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686K08164 (from clone DKFZp686K08164).
ACCESSION CR749316
VERSION CR749316.1 GI:51476361
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 7777)
AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Oeanger, A., Fobo, G., Han, M. and Wiemann, S.
CONSRPTM The German CDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMPZ (Biomedical Research Center at the Heinrich-Heine-University, Duesseidorf/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZp686K08164) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering.
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686K08164
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
1..7777
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686K08164"
/tissue types="uterus endothel, primary cell culture"
/clone lib="686 (synonym: hlcc3). Vector pSport1_SfiI; host DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="fibronectin 1, differentially spliced"
1..7777
/gene="DKFZp686K08164"
266..7069
/gene="DKFZp686K08164"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18171.1"

[illegible]

ORIGIN

Query Match	87.8%	Score 1885.8	DB 3	Length 7777
Best Local Similarity	97.4%	Pred. No. 0		
Matches 1945	Conservative 0	Mismatches 7	Indels 45	Gaps 1
Qy 1	CAAACTTGGTGGCAACTTGCTCCCGGTGGGGGCTCTCCCCACCGGTCTCAACATGC	60		
Db 210	CAAACTTGGTGGCAACTTGCTCCCGGTGGGGGCTCTCCCCACCGGTCTCAACATGC	269		
Qy 61	TTAGGGGTTCGGGGGCCCGGGCTGCTGCTGTGGCGGTCCAGTGCCTGGGACACAGCGGTGC	120		
Db 270	TTAGGGGTTCGGGGGCCCGGGCTGCTGCTGTGGCGGTCTGTGCTGGGACACAGCGGTGC	329		
Qy 121	CCTCCACGGGAGCCTCGAAGACGAAGGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCC	180		
Db 330	CCTCCACGGGAGCCTCGAAGACGAAGGCGGGCTCAGCAAAATGGTTTCAGCCCGAGTCCC	389		
Qy 181	CGTGGCTGTGAGTCAAAAGCAAGCCCGGTTGTTATGACAATGGAAAAACATATCAGATAA	240		
Db 390	CGTGGCTGTGAGTCAAAAGCAAGCCCGGTTGTTATGACAATGGAAAAACATATCAGATAA	449		
Qy 241	ATCAACAGTGGGAGCGGACCTTACCTAGGCAATGGTTGGTTGACCTTGTATGGAGGAA	300		
Db 450	ATCAACAGTGGGAGCGGACCTTACCTAGGCAATGGTTGGTTGACCTTGTATGGAGGAA	509		
Qy 301	GCCGAGCTTTTAACTGCGAGAGTAAAACCTGAAGCTGAAGACCTTGCTTTGACAAAGTACA	360		
Db 510	GCCGAGCTTTTAACTGCGAGAGTAAAACCTGAAGCTGAAGACCTTGCTTTGACAAAGTACA	569		
Qy 361	CTGGGAACACTTACCGGAGTGGTGACACTTATGAGCGGTCTTAAAGACTCCATGATCTGGG	420		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2005, 01:52:51 ; Search time 1727 Seconds
(without alignments)
10281.009 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaacttggtggaacttgc.....aaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2120.6	98.8	2127	22	US-10-956-157-4288
2	2119.8	98.7	2443	20	US-10-741-601-70
3	2119.8	98.7	2443	22	US-10-741-600-238
4	2119	98.7	2127	17	US-10-210-120-49
5	2119	98.7	2127	24	US-10-909-035-49
6	2119	98.7	2127	24	US-10-909-035-49
7	2119	98.7	2127	24	US-10-909-035-49
8	2119	98.7	2127	24	US-10-909-035-49
9	2119	98.7	2127	24	US-10-909-035-49
10	2119	98.7	2127	24	US-10-909-035-49
11	2119	98.7	2127	24	US-10-909-035-49
12	2119	98.7	2127	24	US-10-909-035-49
13	2119	98.7	2127	24	US-10-909-035-49
14	2119	98.7	2127	24	US-10-909-035-49
15	2119	98.7	2127	24	US-10-909-035-49
16	2119	98.7	2127	24	US-10-909-035-49
17	2119	98.7	2127	24	US-10-909-035-49
18	2119	98.7	2127	24	US-10-909-035-49
19	2119	98.7	2127	24	US-10-909-035-49
20	2119	98.7	2127	24	US-10-909-035-49
21	2119	98.7	2127	24	US-10-909-035-49
22	2119	98.7	2127	24	US-10-909-035-49
23	2119	98.7	2127	24	US-10-909-035-49
24	2119	98.7	2127	24	US-10-909-035-49
25	2119	98.7	2127	24	US-10-909-035-49
26	2119	98.7	2127	24	US-10-909-035-49
27	2119	98.7	2127	24	US-10-909-035-49
28	2119	98.7	2127	24	US-10-909-035-49
29	2119	98.7	2127	24	US-10-909-035-49
30	2119	98.7	2127	24	US-10-909-035-49
31	2119	98.7	2127	24	US-10-909-035-49
32	2119	98.7	2127	24	US-10-909-035-49
33	2119	98.7	2127	24	US-10-909-035-49
34	2119	98.7	2127	24	US-10-909-035-49
35	2119	98.7	2127	24	US-10-909-035-49
36	2119	98.7	2127	24	US-10-909-035-49
37	2119	98.7	2127	24	US-10-909-035-49
38	2119	98.7	2127	24	US-10-909-035-49
39	2119	98.7	2127	24	US-10-909-035-49
40	2119	98.7	2127	24	US-10-909-035-49
41	2119	98.7	2127	24	US-10-909-035-49
42	2119	98.7	2127	24	US-10-909-035-49
43	2119	98.7	2127	24	US-10-909-035-49
44	2119	98.7	2127	24	US-10-909-035-49
45	2119	98.7	2127	24	US-10-909-035-49

6	2064.8	96.2	2488	20	US-10-741-601-75	Sequence 75, Appl
7	2064.8	96.2	2488	22	US-10-741-600-244	Sequence 244, Appl
8	2064.8	96.2	2488	22	US-10-741-600-244	Sequence 244, Appl
9	1975.8	92.0	4295	18	US-10-144-194A-51	Sequence 51, Appl
10	1975.8	92.0	4295	18	US-10-491-566-51	Sequence 51, Appl
11	1890.6	88.1	8027	18	US-10-447-161-8	Sequence 8, Appl
12	1890.6	88.1	8027	20	US-10-734-564-27	Sequence 27, Appl
13	1890.6	88.1	8027	22	US-10-852-335A-53	Sequence 53, Appl
14	1890.6	88.1	8027	24	US-10-287-436A-81	Sequence 81, Appl
15	1890.6	88.1	8815	22	US-10-868-577A-62	Sequence 62, Appl
16	1890.6	88.1	8815	22	US-10-868-549-21	Sequence 21, Appl
17	1889.8	88.0	6510	20	US-10-741-601-72	Sequence 72, Appl
18	1889.8	88.0	6510	20	US-10-741-600-241	Sequence 241, Appl
19	1889.8	88.0	7823	20	US-10-741-601-77	Sequence 77, Appl
20	1889.8	88.0	7823	22	US-10-741-600-245	Sequence 245, Appl
21	1889.8	88.0	7848	20	US-10-741-601-78	Sequence 78, Appl
22	1889.8	88.0	7935	20	US-10-741-600-246	Sequence 246, Appl
23	1889.8	88.0	7935	22	US-10-741-601-74	Sequence 74, Appl
24	1889.8	88.0	7935	22	US-10-741-600-240	Sequence 240, Appl
25	1889.8	88.0	7959	20	US-10-741-601-81	Sequence 81, Appl
26	1889.8	88.0	7959	22	US-10-741-600-249	Sequence 249, Appl
27	1889.8	88.0	8013	20	US-10-741-601-71	Sequence 71, Appl
28	1889.8	88.0	8013	22	US-10-741-600-242	Sequence 242, Appl
29	1889.8	88.0	8155	20	US-10-741-601-79	Sequence 79, Appl
30	1889.8	88.0	8155	22	US-10-741-600-247	Sequence 247, Appl
31	1889.8	88.0	8226	20	US-10-741-601-69	Sequence 69, Appl
32	1889.8	88.0	8226	22	US-10-741-600-237	Sequence 237, Appl
33	1889.8	88.0	8278	20	US-10-741-601-82	Sequence 82, Appl
34	1889.8	88.0	8278	22	US-10-741-600-250	Sequence 250, Appl
35	1889.8	88.0	8332	20	US-10-741-601-73	Sequence 73, Appl
36	1889.8	88.0	8332	22	US-10-741-600-239	Sequence 239, Appl
37	1889.8	88.0	8371	20	US-10-741-601-76	Sequence 76, Appl
38	1889.8	88.0	8371	22	US-10-741-600-243	Sequence 243, Appl
39	1887.4	87.9	7795	16	US-10-084-817-2	Sequence 2, Appl
40	1887.4	87.9	7867	14	US-10-098-841-6	Sequence 6, Appl
41	1887.4	87.9	8044	17	US-10-240-965-121	Sequence 121, Appl
42	1887.4	87.9	8044	24	US-10-765-700-135	Sequence 135, Appl
43	1887.4	87.9	8062	14	US-10-098-841-5	Sequence 5, Appl
44	1887.4	87.9	8137	14	US-10-098-841-8	Sequence 8, Appl
45	1839.6	85.7	7361	19	US-10-236-392-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-956-157-4288
; Sequence 4288, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4288
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-4288

Query Match 98.8%; Score 2120.6; DB 22; Length 2127;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2123; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 CAAACTTGGTGGCAACTTGGCTCCCGGTGGGGGGTCTCTCCCGCACCGTCTCAACATGC 60
Db 1 CAAACTTGGTGGCAACTTGGCTCCCGGTGGGGGGTCTCTCTCCCGCACCGTCTCAACATGC 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2005, 23:37:45 ; Search time 237 Seconds
(without alignments)

2819.365 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 2147

Sequence: 1 caaactgttggaacttgc.....aaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New.*

- 1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
- 10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	1890.6	88.1	8232	6 US-10-821-234-693	Sequence 693, App
2	772.2	36.0	777	6 US-10-631-558-3	Sequence 3, Appli
3	330	15.4	1541	6 US-10-631-558-9	Sequence 9, Appli
4	330	15.4	2096	6 US-10-631-558-12	Sequence 12, Appl
5	312	14.5	1661	6 US-10-631-558-10	Sequence 10, Appl
6	232.8	10.8	1782	6 US-10-631-558-11	Sequence 11, Appl
7	162.2	7.6	2154	6 US-10-750-185-25146	Sequence 25146, A
8	73.8	3.4	2334	8 US-11-112-944-22	Sequence 22, Appl
9	71.8	3.3	3069	6 US-10-909-125-824	Sequence 824, App
c 10	71.8	3.3	3515	6 US-10-821-234-65	Sequence 65, Appl
11	48.2	2.2	1672	7 US-11-102-240-17	Sequence 17, Appl
12	37.6	1.8	257	6 US-10-623-155-48	Sequence 48, Appl
13	35.4	1.6	2822	6 US-10-131-826A-305	Sequence 305, App
14	34.8	1.6	1113	6 US-10-750-185-41675	Sequence 41675, A
15	34.8	1.6	2140	7 US-11-087-227-17	Sequence 17, Appl
16	34.8	1.6	2281	7 US-11-087-227-19	Sequence 19, Appl
c 17	34.8	1.6	149419	7 US-11-112-908-49	Sequence 49, Appl
c 18	34.8	1.6	166111	7 US-11-112-908-47	Sequence 47, Appl
19	34.6	1.6	1511	7 US-11-102-978-6	Sequence 6, Appli
20	34.6	1.6	4053	6 US-10-131-826A-351	Sequence 351, App
c 21	34.6	1.6	319608	7 US-11-145-703-1	Sequence 1, Appli
22	34.4	1.6	2745	6 US-10-517-544-76	Sequence 76, Appl
c 23	34.2	1.6	1828	6 US-10-510-386-99	Sequence 99, Appl

ALIGNMENTS

RESULT 1

US-10-821-234-693/c
; Sequence 693, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821.234
; CURRENT FILING DATE: 2004-04-07
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 693
; LENGTH: 8232
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-821-234-693

Query Match	88.1%	Score 1890.6;	DB 6;	Length 8232;
Best Local Similarity	97.5%;	Pred. No. 0;		
Matches 1948;	Conservative	0;	Mismatches	4;
Indels	45;	Gaps	1;	
Qy	1	CAAACTTGTGGCAACTTGGCTCCCGTGGCGGCTCTCCCGCCACCGTCTCAACATGC	60	
Db	8021	CAAACTTGTGGCAACTTGGCTCCCGTGGCGGCTCTCTCCCGCCACCGTCTCAACATGC	7962	
Qy	61	TTAGGGTCCGGGCGCGGCTGCTGCTGCTGCGCGTCCAGTCCAGTGGGACACGGTGC	120	
Db	7961	TTAGGGTCCGGGCGCGGCTGCTGCTGCTGCGCGTCCAGTGGGACACGGTGC	7902	
Qy	121	CTCCACGGAGCTCGAAGCAAGAGGAGGCTCAGCAATGGTTTTCAGCCCCAGTCCC	180	
Db	7901	CTCCACGGAGCTCGAAGCAAGAGGAGGCTCAGCAATGGTTTTCAGCCCCAGTCCC	7842	
Qy	181	CGGTGGTGTGAGTCAAGCAAGCCCGGTTTATGACAATGGAAACACTATCAGATAA	240	
Db	7841	CGGTGGTGTGAGTCAAGCAAGCCCGGTTTATGACAATGGAAACACTATCAGATAA	7782	
Qy	241	ATCAACAGTGGGACCGGACCTCTAGGCAATGGTTTGTACTTGTATGGAGAA	300	
Db	7781	ATCAACAGTGGGACCGGACCTCTAGGCAATGGTTTGTACTTGTATGGAGAA	7722	

Sequence 43048, A
Sequence 58334, A
Sequence 1, Appli
Sequence 31, Appli
Sequence 830, App
Sequence 4450, Ap
Sequence 25822, A
Sequence 36396, A
Sequence 43012, A
Sequence 31441, A
Sequence 25004, A
Sequence 34742, A
Sequence 39432, A
Sequence 42, Appl
Sequence 41, Appl
Sequence 54838, A
Sequence 51832, A
Sequence 52709, A
Sequence 24994, A
Sequence 63106, A

c 24 34 1.6 1916 6 US-10-750-185-43048
25 34 1.6 2001 6 US-10-750-185-58334
26 33.8 1.6 1258 6 US-10-997-437A-1
27 33.8 1.6 2653 7 US-11-073-347-2
28 33.8 1.6 207600 7 US-11-112-908-31
29 33.6 1.6 1358 6 US-10-909-125-830
30 33.4 1.6 3699 6 US-10-793-626-4450
c 31 33.4 1.6 3980 6 US-10-750-185-25822
c 32 33.2 1.5 802 6 US-10-750-185-36396
c 33 33 1.5 698 6 US-10-750-185-43012
c 34 33 1.5 997 6 US-10-750-185-31441
c 35 33 1.5 1717 6 US-10-750-185-25004
c 36 32.8 1.5 1623 6 US-10-750-185-34742
c 37 32.8 1.5 4233 6 US-10-750-185-39432
c 38 32.8 1.5 15515 7 US-11-112-908-42
c 39 32.8 1.5 159660 7 US-11-112-908-43
c 40 32.8 1.5 177623 7 US-11-112-908-41
c 41 32.6 1.5 858 6 US-10-750-185-54838
c 42 32.6 1.5 945 6 US-10-750-185-51832
c 43 32.6 1.5 1245 6 US-10-750-185-52709
c 44 32.4 1.5 1114 6 US-10-750-185-24994
c 45 32.4 1.5 1181 6 US-10-750-185-63106

Qy 301 GCCGAGGTTTTAACTGCGAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACA 360
 Db 7721 GCCGAGGTTTTAACTGCGAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACA 7662
 Qy 361 CTGGGAACACTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTTCCATGATCTGGG 420
 Db 7661 CTGGGAACACTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTTCCATGATCTGGG 7602
 Qy 421 ACTGTACTGATCGGGCTGGGCGAGGGAGAAATAAGCTGTACCATCGCAACCGCTGCC 480
 Db 7601 ACTGTACTGATCGGGCTGGGCGAGGGAGAAATAAGCTGTACCATCGCAACCGCTGCC 7542
 Qy 481 ATGAAGGGGGTCACTCTTACAAGATTGGTGACACTGGAGGAGACCATGAGACTGGTG 540
 Db 7541 ATGAAGGGGGTCACTCTTACAAGATTGGTGACACTGGAGGAGACCATGAGACTGGTG 7482
 Qy 541 GTTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGAGGAATGGAACCTGCAAGCCCA 600
 Db 7481 GTTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGAGGAATGGAACCTGCAAGCCCA 7422
 Qy 601 TAGCTGAGAAGTGTGATCATGCTGTGGACTTCTCTATGTGGTGGAGAAACGTGGG 660
 Db 7421 TAGCTGAGAAGTGTGATCATGCTGTGGACTTCTCTATGTGGTGGAGAAACGTGGG 7362
 Qy 661 AGAAGCCCTACCAAGGCTGGATGATGGTGAATTTGATCTGGGAGAGCGAGCGAC 720
 Db 7361 AGAAGCCCTACCAAGGCTGGATGATGGTGAATTTGATCTGGGAGAGCGAGCGAC 7302
 Qy 721 GCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTATAGAA 780
 Db 7301 GCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTATAGAA 7242
 Qy 781 TTGAGACACCTGAGCAAGAGGATTAATCAGGAAACCTGCTCAGTGCATCTGCACAG 840
 Db 7241 TTGAGACACCTGAGCAAGAGGATTAATCAGGAAACCTGCTCAGTGCATCTGCACAG 7182
 Qy 841 GCAACGGCCGAGGAGTGGAAAGTGTGAGAGGCAACCTCTGTGAGACCAACATCGAGCG 900
 Db 7181 GCAACGGCCGAGGAGTGGAAAGTGTGAGAGGCAACCTCTGTGAGACCAACATCGAGCG 7122
 Qy 901 GATCTGGCCCTTCAACCGATGTTGTCAGAGTGTGTTTACCAACCGAGCTTCAACCCCAAGC 960
 Db 7121 GATCTGGCCCTTCAACCGATGTTGTCAGAGTGTGTTTACCAACCGAGCTTCAACCCCAAGC 7062
 Qy 961 CTCCTCCCTATGGCCACTGTGTCAAGACAGTGGTGTGCTTCTGTGGGATGCAGT 1020
 Db 7061 CTCCTCCCTATGGCCACTGTGTCAAGACAGTGGTGTGCTTCTGTGGGATGCAGT 7002
 Qy 1021 GGCTGAAGACACAAAGGAATTAAGCAAAATGCTTTCACGCTGCTGGCAACCGAGTCACT 1080
 Db 7001 GGCTGAAGACACAAAGGAATTAAGCAAAATGCTTTCACGCTGCTGGCAACCGAGTCACT 6942
 Qy 1081 GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTGTCT 1140
 Db 6941 GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTGTCT 6882
 Qy 1141 TACATTTACCTAACAAGCAGGAC----- 1165
 Db 6881 TACATTTACCTAACAAGCAGGAC----- 6822
 Qy 1166 -----GGACAGACAACTTCGAATTTATGACGAGGACAGAAATACTCTTTCTGCA 1215
 Db 6821 GACATCTTTGGTGACGACAACTTCGAATTTATGACGAGGACAGAAATACTCTTTCTGCA 6762
 Qy 1216 CAGACACACTGTTTTGGTTTCAGACTCGAGGAGGAAATTTCCAAATGGTGCCTTTGTGCCACT 1275
 Db 6761 CAGACACACTGTTTTGGTTTCAGACTCGAGGAGGAAATTTCCAAATGGTGCCTTTGTGCCACT 6702
 Qy 1276 TCCCTCTTCTATAACAACCAACCAATTTACATGATTGCACTTCTGAGGGCAGAGAGACA 1335
 Db 6701 TCCCTCTTCTATAACAACCAACCAATTTACATGATTGCACTTCTGAGGGCAGAGAGACA 6642

Qy 1336 ACATGAAGTGGTGTGGGACCAACAGAACTATGATCCGACCAAGAGTGTGGGTTCGCC 1395
 Db 6641 ACATGAAGTGGTGTGGGACCAACAGAACTATGATCCGACCAAGAGTGTGGGTTCGCC 6582
 Qy 1396 CCATGGTGGCCACGAGGAAATCTGCACAAACCAATGAAGGGGTCACTACCGCATTTGGAG 1455
 Db 6581 CCATGGTGGCCACGAGGAAATCTGCACAAACCAATGAAGGGGTCACTACCGCATTTGGAG 6522
 Qy 1456 ATCAGTGGGATAAGCAGCATGACATGGTGCACATGATGAGTGCACGTGTGTGGGAATG 1515
 Db 6521 ATCAGTGGGATAAGCAGCATGACATGGTGCACATGATGAGTGCACGTGTGTGGGAATG 6462
 Qy 1516 GTCGTGGGGAATGGACATGCAATTCCTTACTCGCAGCTTCGAGATCAGTGCATTTGATG 1575
 Db 6461 GTCGTGGGGAATGGACATGCAATTCCTTACTCGCAGCTTCGAGATCAGTGCATTTGATG 6402
 Qy 1576 ACATCACTTCAATGTGAACGACACATTCACCAAGCGTCAATGAAGGGGCGACATGCTGA 1635
 Db 6401 ACATCACTTCAATGTGAACGACACATTCACCAAGCGTCAATGAAGGGGCGACATGCTGA 6342
 Qy 1636 ACTGTACATCTTTCGGTTCAGGGTCGGGCGAGGTGGAAGTGTGATCCCGTGCACCAATGCC 1695
 Db 6341 ACTGTACATCTTTCGGTTCAGGGTCGGGCGAGGTGGAAGTGTGATCCCGTGCACCAATGCC 6282
 Qy 1696 AGGATTCAGAGACTGGGACGTTTTTATCAAAATGGAGATTTCATGGGAGAAATGATGTCATG 1755
 Db 6281 AGGATTCAGAGACTGGGACGTTTTTATCAAAATGGAGATTTCATGGGAGAAATGATGTCATG 6222
 Qy 1756 GTGTGAGATACAGTGTCTACTGTATGGCCGTGGCATTTGGGAGTGGCATTTGCCAACCTT 1815
 Db 6221 GTGTGAGATACAGTGTCTACTGTATGGCCGTGGCATTTGGGAGTGGCATTTGCCAACCTT 6162
 Qy 1816 TACAGACTTATCAAGCTCAAGTGGTCTCTGCAAGTATTTATCAGTGAAGTCCGAGTCCGAGTC 1875
 Db 6161 TACAGACTTATCAAGCTCAAGTGGTCTCTGCAAGTATTTATCAGTGAAGTCCGAGTCCGAGTC 6102
 Qy 1876 AGCCCACTCCCAACCCCATCCAGTGGAAATGACCAAGCCATCTCACATTTCCAAAGTACA 1935
 Db 6101 AGCCCACTCCCAACCCCATCCAGTGGAAATGACCAAGCCATCTCACATTTCCAAAGTACA 6042
 Qy 1936 TTCTCAGGTGGAGACCT 1952
 Db 6041 TTCTCAGGTGGAGACCT 6025

RESULT 2
 US-10-631-558-3
 ; Sequence 3, Application US/10631558
 ; Publication No. US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Rajagopal, Kammar
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/10/631,558
 ; CURRENT FILING DATE: 2003-07-31
 ; PRIOR APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2005, 21:28:04 ; Search time 1571 Seconds
(without alignments)
11301.308 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaactgtgtggaacttgc.....aaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2120.6	98.8	2127	9	US-10-956-157-4288
2	2119.8	98.7	2443	7	US-10-741-601-70
3	2119.8	98.7	2443	8	US-10-741-600-238
4	2119	98.7	2127	6	US-10-210-320-49
5	2119	98.7	2127	9	US-10-909-035-49
6	2064.8	96.2	2488	7	US-10-741-601-75
7	2064.8	96.2	2488	8	US-10-741-600-244
8	1975.8	92.0	4295	6	US-10-144-194A-51
9	1975.8	92.0	4295	8	US-10-491-566-51
10	1890.6	88.1	8027	7	US-10-447-161-8
11	1890.6	88.1	8027	7	US-10-734-564-27
12	1890.6	88.1	8027	9	US-10-852-335A-53
13	1890.6	88.1	8027	9	US-10-287-436A-81
14	1890.6	88.1	8815	8	US-10-868-577A-62
15	1890.6	88.1	8815	8	US-10-868-549-21
16	1889.8	88.0	6510	7	US-10-741-601-72
17	1889.8	88.0	6510	8	US-10-741-600-241
18	1889.8	88.0	7823	7	US-10-741-601-77
19	1889.8	88.0	7823	8	US-10-741-600-245
20	1889.8	88.0	7848	7	US-10-741-601-78
21	1889.8	88.0	7848	8	US-10-741-600-246
22	1889.8	88.0	7935	7	US-10-741-601-74
23	1889.8	88.0	7935	8	US-10-741-600-240

24 1889.8 88.0 7959 7 US-10-741-601-81 Sequence 81, Appl
25 1889.8 88.0 7959 8 US-10-741-600-249 Sequence 249, Appl
26 1889.8 88.0 8013 7 US-10-741-601-71 Sequence 71, Appl
27 1889.8 88.0 8013 8 US-10-741-600-242 Sequence 242, Appl
28 1889.8 88.0 8155 7 US-10-741-601-79 Sequence 79, Appl
29 1889.8 88.0 8155 8 US-10-741-600-247 Sequence 247, Appl
30 1889.8 88.0 8226 7 US-10-741-601-69 Sequence 69, Appl
31 1889.8 88.0 8226 8 US-10-741-600-237 Sequence 237, Appl
32 1889.8 88.0 8278 7 US-10-741-601-82 Sequence 82, Appl
33 1889.8 88.0 8278 8 US-10-741-600-250 Sequence 250, Appl
34 1889.8 88.0 8332 7 US-10-741-601-73 Sequence 73, Appl
35 1889.8 88.0 8332 8 US-10-741-600-239 Sequence 239, Appl
36 1889.8 88.0 8371 7 US-10-741-601-76 Sequence 76, Appl
37 1889.8 88.0 8371 8 US-10-741-600-243 Sequence 243, Appl
38 1887.4 87.9 7795 5 US-10-084-817-2 Sequence 2, Appl
39 1887.4 87.9 7867 5 US-10-098-841-6 Sequence 6, Appl
40 1887.4 87.9 8044 6 US-10-240-965-121 Sequence 121, Appl
41 1887.4 87.9 8044 9 US-10-765-700-135 Sequence 135, Appl
42 1887.4 87.9 8062 5 US-10-098-841-5 Sequence 5, Appl
43 1887.4 87.9 8137 5 US-10-098-841-8 Sequence 8, Appl
44 1887.4 87.9 8230 5 US-10-098-841-7 Sequence 7, Appl
45 1839.6 85.7 7361 7 US-10-236-392-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-956-157-4288

; Sequence 4288, Application US/10956157

; Publication No. US20050118625A1

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Mounts, William

; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES

; FILE REFERENCE: 031896-043000 (AM 101081)

; CURRENT APPLICATION NUMBER: US/10/956,157

; CURRENT FILING DATE: 2004-10-04

; NUMBER OF SEQ ID NOS: 319805

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 4288

; LENGTH: 2127

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-10-956-157-4288

Query Match 98.8%; Score 2120.6; DB 9; Length 2127;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2123; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CAAACTTGGTGGCAACTTGGCTCCCGTGGGGGGTCTCTCCCCACCGTCTCAACATGC 60

Db 1 CAAACTTGGTGGCAACTTGGCTCCCGTGGGGGGTCTCTCCCCACCGTCTCAACATGC 60

Qy 61 TTAGGGGTCCGGGGCCGGGCTGCTGCTGTGCGCCGTCAGTGTCTGGGACAGCGGTGC 120

Db 61 TTAGGGGTCCGGGGCCGGGCTGCTGCTGTGCGCCGTCAGTGTCTGGGACAGCGGTGC 120

Qy 121 CTTCACGGGAGCTCGAAGCAAGGAGGAGGCTCAGCAATGGTTTCAGCCCCAGTCCC 180

Db 121 CTTCACGGGAGCTCGAAGCAAGGAGGAGGCTCAGCAATGGTTTCAGCCCCAGTCCC 180

Qy 181 CGGTGGCTGTGAGTCAAGCAAGCCCGTTGTTATGCAATGGAACACATATCAGATAA 240

Db 181 CGGTGGCTGTGAGTCAAGCAAGCCCGTTGTTATGCAATGGAACACATATCAGATAA 240

Qy 241 ATCAACAGTCGGGACCGGACCTAGCGCAATCGTTGGTTGTACTTGTATGAGGAA 300

Db 241 ATCAACAGTCGGGACCGGACCTAGCGCAATCGTTGGTTGTACTTGTATGAGGAA 300

Qy 301 GCCGAGGTTTAACTGCGAGAGTAAACCTGAAGCTGAAGACTTGTCTTGACAAGTACA 360

|||||


```
Db 301 GCCAGGTTTTAACTGCGAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAGTACA 360
Qy 361 CTGGGAACACTTACCGAGTGGGTGACACTTATGAGCGTCCTTAAAGACTTCCATGATCTGGG 420
Db 361 CTGGGAACACTTACCGAGTGGGTGACACTTATGAGCGTCCTTAAAGACTTCCATGATCTGGG 420
Qy 421 ACTGTACTCTGCATCGGGCTGGGGCGAGGGAATAAGCTGTACCATCGCAACCGCTGCC 480
Db 421 ACTGTACTCTGCATCGGGCTGGGGCGAGGGAATAAGCTGTACCATCGCAACCGCTGCC 480
Qy 481 ATGAAGGGGTGAGTCTCAAGAATTGGTGAACCTGGAGGAGACCATGAGACTGGTG 540
Db 481 ATGAAGGGGTGAGTCTCAAGAATTGGTGAACCTGGAGGAGACCATGAGACTGGTG 540
Qy 541 GTTACATGTTAGAGTGTGTCTTGTGTAATGAAAGGAGAAATGGAACCTGCAAGGCCA 600
Db 541 GTTACATGTTAGAGTGTGTCTTGTGTAATGAAAGGAGAAATGGAACCTGCAAGGCCA 600
Qy 601 TAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCCATGTGGTGGGAAACGTTGGG 660
Db 601 TAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCCATGTGGTGGGAAACGTTGGG 660
Qy 661 AGAAGCCCTACCAAGGCTGGATGATGTGATGATGTAATTGCTGGCTGGGAGAGGCGAGAC 720
Db 661 AGAAGCCCTACCAAGGCTGGATGATGTGATGATGTAATTGCTGGCTGGGAGAGGCGAGAC 720
Qy 721 GCATCATTGACCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTATAGAA 780
Db 721 GCATCATTGACCTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTATAGAA 780
Qy 781 TTGGAGACACTGGAGCAAGAGGATATCAGGAAACCTGCTCCAGTGCATCTGCACAG 840
Db 781 TTGGAGACACTGGAGCAAGAGGATATCAGGAAACCTGCTCCAGTGCATCTGCACAG 840
Qy 841 GCAAGCGCGAGGAGAGTGGAAAGTGTGAGAGGACACCTCTGTGAGACACATCGAGCG 900
Db 841 GCAAGCGCGAGGAGAGTGGAAAGTGTGAGAGGACACCTCTGTGAGACACATCGAGCG 900
Qy 901 GATCTGCGCCCTTACACGATGCTGTCGAGCTGTTTACCAACCGCAGCTCACCCCGAGC 960
Db 901 GATCTGCGCCCTTACACGATGCTGTCGAGCTGTTTACCAACCGCAGCTCACCCCGAGC 960
Qy 961 CTCCTCCCTATGGCCACTGTCTCAGACAGTGGTGTCTACTCTGTGGGGATGCAGT 1020
Db 961 CTCCTCCCTATGGCCACTGTCTCAGACAGTGGTGTCTACTCTGTGGGGATGCAGT 1020
Qy 1021 GGCTGAAGACACAAGGAAATAGGAAATGCTTTGCA CGTGCCTGGGCAACGGAGTCAGCT 1080
Db 1021 GGCTGAAGACACAAGGAAATAGGAAATGCTTTGCA CGTGCCTGGGCAACGGAGTCAGCT 1080
Qy 1081 GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCACTCAAAATGGAGAGCCATGTGTCT 1140
Db 1081 GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCACTCAAAATGGAGAGCCATGTGTCT 1140
Qy 1141 TACCATTCCCTACACGACAGGAGCGGACAGCAACACTTTCGAAATTTATGAGCAGGACCAGA 1200
Db 1141 TACCATTCCCTACACGAGGAGCGTGCAGCAACACTTCGAAATTTATGAGCAGGACCAGA 1200
Qy 1201 AATACTCTTTTGCACAGACACACTGTTTGGTTTGCAGACTCGAGGAGGAAATTCCTCATG 1260
Db 1201 AATACTCTTTTGCACAGACACACTGTTTGGTTTGCAGACTCGAGGAGGAAATTCCTCATG 1260
Qy 1261 GTGCTTGTGCGACTTCCCTTCTATACACAAACCAATTTACACTGATTTGCACTTCTG 1320
Db 1261 GTGCTTGTGCGACTTCCCTTCTATACAAACCAATTTACACTGATTTGCACTTCTG 1320
Qy 1321 AGGGCAGAGAGACAACTAAGTGGTGGTGGGACCAACAGAACTATGATGCCGACCAAGA 1380
Db 1321 AGGGCAGAGAGACAACTAAGTGGTGGTGGGACCAACAGAACTATGATGCCGACCAAGA 1380
Qy 1381 AGTTTGGGTTCTGCCCATGGCTGCCACGAGGAAATCTGCACACCAATGAAGGGGTCA 1440
Db 1381 AGTTTGGGTTCTGCCCATGGCTGCCACGAGGAAATCTGCACACCAATGAAGGGGTCA 1440
```

```
Qy 1441 TGTACCGCATTTGGAGATCAGTGGGATAAGCAGCATGACATGGGTACATGATGAGGTGCA 1500
Db 1441 TGTACCGCATTTGGAGATCAGTGGGATAAGCAGCATGACATGGGTACATGATGAGGTGCA 1500
Qy 1501 CGTGTGTTGGGAATGTCGTGGGGAATGGACATGCATTCCTACTCGAGCTTCGAGATC 1560
Db 1501 CGTGTGTTGGGAATGTCGTGGGGAATGGACATGCATTCCTACTCGAGCTTCGAGATC 1560
Qy 1561 AGTGCAATGTTGATGACATCACTTACAAATGTGAACGACACATTCACAAGCGTCATGAAG 1620
Db 1561 AGTGCAATGTTGATGACATCACTTACAAATGTGAACGACACATTCACAAGCGTCATGAAG 1620
Qy 1621 AGGGGCACATGCTGGAATGCTGCTTCGGTTCAGGCTCGGGGAGGTGGAAAGTGTGATC 1680
Db 1621 AGGGGCACATGCTGGAATGCTGCTTCGGTTCAGGCTCGGGGAGGTGGAAAGTGTGATC 1680
Qy 1681 CCGTCGACCAATGCCAGGATTCAGAGACTCGGACCTGCTATGGCCGTGGCAATTCATGGG 1740
Db 1681 CCGTCGACCAATGCCAGGATTCAGAGACTCGGACCTGCTATGGCCGTGGCAATTCATGGG 1740
Qy 1741 AGAAGTATGTGATGCTGCTGAGATACCAAGTCTGCTATGGCCGTGGCAATTCATGGG 1800
Db 1741 AGAAGTATGTGATGCTGCTGAGATACCAAGTCTGCTATGGCCGTGGCAATTCATGGG 1800
Qy 1801 GGCAATTCGCAACCTTTTACAGACCTTATCCAAAGCTCAAGTGGTCTGTGCTGAAAGTATTTATCA 1860
Db 1801 GGCAATTCGCAACCTTTTACAGACCTTATCCAAAGCTCAAGTGGTCTGTGCTGAAAGTATTTATCA 1860
Qy 1861 CTGAGACTCGGAGTCAGCCCAACTCCACCCCATCCAGTGGAAATGCCACCAAGCATCTC 1920
Db 1861 CTGAGACTCGGAGTCAGCCCAACTCCACCCCATCCAGTGGAAATGCCACCAAGCATCTC 1920
Qy 1921 ACATTTCCAAAGTACATTTCTCAGGTGGAGACTGTGAGTATCCCAACCCAGAAACCTTGGAT 1980
Db 1921 ACATTTCCAAAGTACATTTCTCAGGTGGAGACTGTGAGTATCCCAACCCAGAAACCTTGGAT 1980
Qy 1981 ACTGAGTCTCCTAATCTTATCAATCTTGATGGTCTCTTTTCCAGCTTTTGAGCCAA 2040
Db 1981 ACTGAGTCTCCTAATCTTATCAATCTTGATGGTCTCTTTTCCAGCTTTTGAGCCAA 2040
Qy 2041 CAATCTGATTAATCTTCTATAGCAATTTACTATATTTGTTAGTGACAAACAATATG 2100
Db 2041 CAATCTGATTAATCTTCTATAGCAATTTACTATATTTGTTAGTGACAAACAATATG 2100
Qy 2101 TGGTCAATTAATTTGACTTTGTAGACTG 2127
Db 2101 TGGTCAATTAATTTGACTTTGTAGACTG 2127
```

RESULT 2

```
US-10-741-601-70
; Sequence 70, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-70
```

```
Query Match 98.7%; Score 2119.8; DB 7; Length 2443;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2121; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2005, 17:53:25 ; Search time 7771 Seconds
(without alignments)
12926.504 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaactgtgtggaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gsa1.*
10: gb_gsa2.*
11: gb_gsa3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1889	88.0	7868	4 HSM806653	BX640608 Homo sapi
2	1889	88.0	8411	4 HSM806992	BX640875 Homo sapi
3	1885.8	87.8	7777	4 CR749316	CR749316 Homo sapi
4	1885.8	87.8	7885	4 CR749281	CR749281 Homo sapi
5	1882.6	87.7	8121	4 CR749317	CR749317 Homo sapi
6	1875.4	87.3	7501	4 BC078656	BC078656 Homo sapi
7	1875.4	87.3	7501	4 BC100030	BC100030 Homo sapi
8	1460.2	68.0	8315	4 AK090135	AK090135 Mus muscu
9	1391.2	64.8	8329	4 AK090130	AK090130 Mus muscu
10	1335.4	62.2	7434	11 DQ039102	DQ039102 Homo sapi
11	1034.6	48.2	7434	11 DQ039103	DQ039103 Pan trogl
C 12	908.6	42.3	1044	5 BX398837	BX398837 BX398837
C 13	875.2	40.8	1052	5 BX380582	BX380582 BX380582
C 14	794.8	37.0	943	5 BX391752	BX391752 BX391752
15	784	36.5	1026	5 BX386270	BX386270 BX386270
16	766.6	35.7	816	1 AU141008	AU141008 AU141008
C 17	766.2	35.7	967	5 BX402381	BX402381 BX402381
18	749	34.9	766	5 CN419594	CN419594 170004245
19	748.8	34.9	1050	5 BX417945	BX417945 BX417945
20	733.2	34.1	765	3 BM715855	BM715855 UI-E-E30-
21	731	34.0	861	1 AU140971	AU140971
22	727.4	33.9	737	1 AU140993	AU140993 AU140993

23	723.2	33.7	731	1	AU140526	AU140526
24	718.2	33.5	739	1	AL706215	AL706215 DKFZp686J
C 25	707.2	32.9	747	6	CA423317	UI-H-PE1-
26	700.2	32.6	729	1	AU140834	AU140834
C 27	693	32.3	695	7	CN419477	170004247
28	693	32.3	706	3	BM674187	UI-E-E30-
29	686.8	32.0	2107	4	AK054456	Mus muscu
30	686.6	32.0	837	1	AU140973	AU140973
31	680.8	31.7	699	1	AU140889	AU140889
32	676.6	31.5	734	1	AU140802	AU140802
33	676	31.5	676	1	AL603362	DKFZp686C
34	673.2	31.4	705	1	AU140789	AU140789
35	669	31.2	723	7	CN419622	170005319
36	664	30.9	667	1	AU140910	AU140910
37	663.2	30.9	911	5	BX398838	BX398838
C 38	662	30.8	738	6	CD613781	55110394J
39	660.6	30.8	670	5	BU620723	UI-H-FL1-
40	656.4	30.6	836	7	CN538822	UI-M-H50-
41	655.4	30.5	1038	5	BX439175	BX439175
42	655.2	30.5	673	1	AU140450	AU140450
43	654	30.5	851	1	AU141024	AU141024
44	652	30.4	859	6	CF616056	AGENCOURT
45	646.4	30.1	796	1	AU140814	AU140814

ALIGNMENTS

RESULT 1	HSM806653	7868 bp	mrna	linear	HTC 20-JAN-2005
LOCUS	HSM806653	Homo sapiens mRNA; cDNA DKFZp686M04163 (from clone DKFZp686M04163).			
DEFINITION	BX640608				
ACCESSION	BX640608.1	GI:34364616			
VERSION	HTC				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 7868)				
AUTHORS	Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C., Oeang, A., Fobo, G., Han, M. and Wiemann, S.				
CONSRM	The German cDNA Consortium				
TITLE	Submitted (20-JAN-2005) MIPS, Ingolstaedter Landstr.1, D-85764				
JOURNAL	Neuherberg, GERMANY				
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686M04163) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M04163 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.				
FEATURES	Location/Qualifiers				
source	1..7868				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="RZPD:DKFZp686M04163Q"				
	/db_xref="taxon:9606"				
	/clone="DKFZp686M04163"				
	/tissue_type="uterus endothel, primary cell culture"				
	/clone_lib="686 (synonym: hlccc3). Vector pSport1_sfi; host DH10B; sites SfiIA + SfiIB"				
	/dev_stage="adult"				
	/note="fibronectin precursor"				
gene	1..7868				
	/gene="DKFZp686M04163"				
CDS	267..7157				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2005, 16:02:29 ; Search time 1207 Seconds
(without alignments)
11855.091 Million cell updates/sec

Title: US-09-581-651d-3
Perfect score: 2147
Sequence: 1 caaactgtggaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2147	100.0	2147	2 AAX81299	Aax81299 Human mig
2	2143.8	99.9	2358	13 ADR67201	Adr67201 Human bla
3	2119.8	98.7	2443	13 ADQ38575	Adq38575 Human SNP
4	2119	98.7	2127	10 ADD18477	Add18477 Human pro
5	2085.6	97.1	2402	14 ADZ26738	Adz26738 Human fib
6	2084.8	96.2	2488	13 ADQ38581	Adq38581 Human SNP
7	1975.8	92.0	4295	8 ACC72037	Acc72037 BC00770 g
8	1925.8	89.7	1929	13 ADS17488	Adsi17488 Nucleotid
9	1925.8	89.7	1929	13 ADR37657	Adr37657 Human fib
10	1925.8	89.7	1929	14 ADW44479	Adw44479 Human fib
11	1925.8	89.7	1929	14 ADY55703	Ady55703 Human fib
12	1925.8	89.7	1929	14 ADY32457	Ady32457 Human fib
13	1925.8	89.7	1929	14 AEB56257	Aeb56257 Svent seq
14	1925.8	89.7	1929	14 AEB78061	Aeb78061 Human DNA
15	1890.6	88.1	7242	13 ADR66637	Adr66637 Human pro
16	1890.6	88.1	7242	13 ADR65953	Adr65953 Human pro
17	1890.6	88.1	7550	8 ACC00412	Acc00412 Human cel
18	1890.6	88.1	7912	14 ADZ26740	Adz26740 Human fib
19	1890.6	88.1	8027	11 ADP64998	Adp64998 Human fib

20	1890.6	88.1	8027	12	ADG89565	Adg89565 Human fib
21	1890.6	88.1	8027	12	ADQ29601	Adq29601 Human col
22	1890.6	88.1	8027	13	ADR67200	Adr67200 Human bla
23	1890.6	88.1	8027	14	ADV70129	Adv70129 Tumor-ass
24	1890.6	88.1	8027	14	AEA04397	Aea04397 Human CDN
25	1890.6	88.1	8027	14	AEB17605	Aeb17605 Human fib
26	1890.6	88.1	8027	14	ADZ26744	Adz26744 Human fib
27	1890.6	88.1	8374	14	ADZ26748	Adz26748 Human fib
28	1890.6	88.1	8449	14	ADZ26564	Adz26564 Human fib
29	1890.6	88.1	8647	14	ADZ26742	Adz26742 Human fib
30	1890.6	88.1	8815	14	ADW81010	Adw81010 Fibronect
31	1890.6	88.1	8815	14	ADX69342	Adx69342 DNA encod
32	1890.6	88.1	8815	14	ADZ26746	Adz26746 Human fib
33	1889.8	88.0	6510	13	ADQ38578	Adq38578 Human SNP
34	1889.8	88.0	7823	13	ADQ38582	Adq38582 Human SNP
35	1889.8	88.0	7848	13	ADQ38583	Adq38583 Human SNP
36	1889.8	88.0	7935	13	ADQ38577	Adq38577 Human SNP
37	1889.8	88.0	7959	13	ADQ38586	Adq38586 Human SNP
38	1889.8	88.0	8013	13	ADQ38579	Adq38579 Human SNP
39	1889.8	88.0	8155	13	ADQ38584	Adq38584 Human SNP
40	1889.8	88.0	8226	13	ADQ38574	Adq38574 Human SNP
41	1889.8	88.0	8278	13	ADQ38587	Adq38587 Human SNP
42	1889.8	88.0	8332	13	ADQ38576	Adq38576 Human SNP
43	1889.8	88.0	8371	13	ADQ38580	Adq38580 Human SNP
44	1887.4	87.9	7795	10	ADJ56196	Adj56196 Zebrafish
45	1887.4	87.9	7867	4	AAI57803	Aai57803 Human poi

ALIGNMENTS

RESULT 1
AAX81299

ID AAX81299 standard; cDNA; 2147 BP.

XX AC AAX81299;

XX DT 21-SEP-1999 (first entry)

XX DE Human migration stimulating factor (MSF) 1-alpha encoding cDNA.

XX KW Migration stimulatory factor; MSF; cell migration; modulation; human;
XX KM wound healing; scarring; MSFI-alpha; ss.

XX OS Homo sapiens.

XX PN WO9931233-A1.

XX PD 24-JUN-1999.

XX PF 15-DEC-1998; 98WO-GB003766.

XX PR 16-DEC-1997; 97GB-00026539.

XX PA (UYDU-) UNIV DUNDEE.

XX PI Schor SL, Schor AM;

XX DR WPI; 1999-430039/36.

XX DR P-PSDB; AAY28901.

XX PT Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX PS Example 1; Fig 1; 86pp; English.

XX CC The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. The present sequence represents the nucleotide sequence encoding a human MSFI-alpha protein

SQL	Sequence	2147 BP; 575 A; 512 C; 575 G; 485 T; 0 U; 0 Other;	
	Query Match	100.0%; Score 2147; DB 2; Length 2147;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 2147; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	CAAACTTGGTGGCAACTTGGCTCCCGGTGCGGGGGTCTCTCCCCACCGTCTCAACATGC	60
Db	1	CAAACTTGGTGGCAACTTGGCTCCCGGTGCGGGGGTCTCTCCCCACCGTCTCAACATGC	60
Qy	61	TTAGGGGTCCGGGGTCCGGGTGCTGCTGCTGCGCGTCCAGTGCCTGGGGACAGCGTGC	120
Db	61	TTAGGGGTCCGGGGTCCGGGTGCTGCTGCTGCGCGTCCAGTGCCTGGGGACAGCGTGC	120
Qy	121	CCTCACGGGAGCCTCGAAGAGCAAGAGGAGGCTCAGCAATGGTTCAGGCCCGAGTCCC	180
Db	121	CCTCACGGGAGCCTCGAAGAGCAAGAGGAGGCTCAGCAATGGTTCAGGCCCGAGTCCC	180
Qy	181	CGGTGGCTGTGAGTCAAAAGCAAGCCCGGTGTTATGACAAATGGAAAAACATATCAGATAA	240
Db	181	CGGTGGCTGTGAGTCAAAAGCAAGCCCGGTGTTATGACAAATGGAAAAACATATCAGATAA	240
Qy	241	ATCAACAGTGGGAGCGGACCTACCTAGGCAATGCGTGGTGTGTTGTTATGGAGGAA	300
Db	241	ATCAACAGTGGGAGCGGACCTACCTAGGCAATGCGTGGTGTGTTGTTATGGAGGAA	300
Qy	301	GCCGAGGTTTAACTGCGAGAGTAAACCTGAGCTGAAGCTGAAGAGACTTGTCTTGCAGATACA	360
Db	301	GCCGAGGTTTAACTGCGAGAGTAAACCTGAGCTGAAGCTGAAGAGACTTGTCTTGCAGATACA	360
Qy	361	CTGGGAAACATTTACCGAGTGGGTGACATTTATGAGCGTCCCTAAAGACTCCCATGATCTGGG	420
Db	361	CTGGGAAACATTTACCGAGTGGGTGACATTTATGAGCGTCCCTAAAGACTCCCATGATCTGGG	420
Qy	421	ACTGTACCTGCATCGGGCTGGGCGAGGGAGAAATGAGCTGTACCATCGCAACCGCTGCC	480
Db	421	ACTGTACCTGCATCGGGCTGGGCGAGGGAGAAATGAGCTGTACCATCGCAACCGCTGCC	480
Qy	481	ATGAGGGGTTCAGTCCCTACAGATTGGTGACACCTGAGGAGACCATGAGACTGGTG	540
Db	481	ATGAGGGGTTCAGTCCCTACAGATTGGTGACACCTGAGGAGACCATGAGACTGGTG	540
Qy	541	GTTACATCTTAGAGTGTGTGTCTTGGTAAATGGAAAAAGGAAATGGACCTGCAAGCCCA	600
Db	541	GTTACATCTTAGAGTGTGTGTCTTGGTAAATGGAAAAAGGAAATGGACCTGCAAGCCCA	600
Qy	601	TAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCTATGTGTGCGGAGAAACGTGGG	660
Db	601	TAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCTATGTGTGCGGAGAAACGTGGG	660
Qy	661	AGAGCCCTACCAAGGCTGGATGATGTTGATGTTGCTGCTGGGAGAGGACCGGAC	720
Db	661	AGAGCCCTACCAAGGCTGGATGATGTTGATGTTGCTGCTGGGAGAGGACCGGAC	720
Qy	721	GCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTCTATAGAA	780
Db	721	GCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTCTATAGAA	780
Qy	781	TTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCAAG	840
Db	781	TTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCAAG	840
Qy	841	GCACGGCCGAGGAGTGGAGTGTGAGAGGACACCTCTGTGCGAGACCATCGAGG	900
Db	841	GCACGGCCGAGGAGTGGAGTGTGAGAGGACACCTCTGTGCGAGACCATCGAGG	900
Qy	901	GATCTGGCCCTTACCGATGTTCTGTGAGCTGTTTACCAACCGCAGGCTCACCCCCAGC	960
Db	901	GATCTGGCCCTTACCGATGTTCTGTGAGCTGTTTACCAACCGCAGGCTCACCCCCAGC	960
Qy	961	CTCTCCCTATGGCCACTGTGTACAGACAGTGGTGTGTCTCTGTGGGATGCAAT	1020
Db	961	CTCTCCCTATGGCCACTGTGTACAGACAGTGGTGTGTCTCTGTGGGATGCAAT	1020

Qy	1021	GGCTGAAGACACAGGAAATAAGCAAAATGCTTTGCAAGTGCCTGGGCAACGGAGTCAGCT	1080
Db	1021	GGCTGAAGACACAGGAAATAAGCAAAATGCTTTGCAAGTGCCTGGGCAACGGAGTCAGCT	1080
Qy	1081	GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCATGTGTCT	1140
Db	1081	GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCATGTGTCT	1140
Qy	1141	TACCAATTCACCTCAACAGCAGGAGCGGACGACCAACTTCGAATTTATGAGCAGGACCCAGA	1200
Db	1141	TACCAATTCACCTCAACAGCAGGAGCGGACGACCAACTTCGAATTTATGAGCAGGACCCAGA	1200
Qy	1201	AATACTCTTTCTGCACAGACCACTGTTTGGTTCAGACTCGAGGAGGAAATTCCTAATG	1260
Db	1201	AATACTCTTTCTGCACAGACCACTGTTTGGTTCAGACTCGAGGAGGAAATTCCTAATG	1260
Qy	1261	GTGCTTGTGGCACTTCCCCTTCTATACAAACCAACTTACACTGATTTGCACTTCTG	1320
Db	1261	GTGCTTGTGGCACTTCCCCTTCTATACAAACCAACTTACACTGATTTGCACTTCTG	1320
Qy	1321	AGGCGAGAGAGCAACATGAAAGTGGTGGGACCAACAGAACTATGATGCCGACCCAGA	1380
Db	1321	AGGCGAGAGAGCAACATGAAAGTGGTGGGACCAACAGAACTATGATGCCGACCCAGA	1380
Qy	1381	AGTTTGGGTTCGCCCCCATGGCTCCCAACGAGGAAATCTGCACAACTGAAGGGGTCA	1440
Db	1381	AGTTTGGGTTCGCCCCCATGGCTCCCAACGAGGAAATCTGCACAACTGAAGGGGTCA	1440
Qy	1441	TGTAACCGCATTTGGAGATCAGTGGGATAGAGCATGAGTGGTCACTGATGAGGTGCA	1500
Db	1441	TGTAACCGCATTTGGAGATCAGTGGGATAGAGCATGAGTGGTCACTGATGAGGTGCA	1500
Qy	1501	CGTGTGTGGGAATGGTGGGAAATGGACATGCACTTGCCTTCTCGAGCTTCGAGATC	1560
Db	1501	CGTGTGTGGGAATGGTGGGAAATGGACATGCACTTGCCTTCTCGAGCTTCGAGATC	1560
Qy	1561	AGTCGATTTGTGATGACATCACTTCAATGTGAACGACACATTCACAAAGCGTTCATGAAG	1620
Db	1561	AGTCGATTTGTGATGACATCACTTCAATGTGAACGACACATTCACAAAGCGTTCATGAAG	1620
Qy	1621	AGGGGCACTGCTGAACTGTATCATGCTTGGTTCAGGCTCGGGGAGGTGGAAGTGTGATC	1680
Db	1621	AGGGGCACTGCTGAACTGTATCATGCTTGGTTCAGGCTCGGGGAGGTGGAAGTGTGATC	1680
Qy	1681	CCGTGCAACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCAAAATGGAGATTCATGGG	1740
Db	1681	CCGTGCAACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCAAAATGGAGATTCATGGG	1740
Qy	1741	AGAAATGATGTCATGGTGTGAGATACCAAGTCTACTGCTATGGCGGTGGGAGT	1800
Db	1741	AGAAATGATGTCATGGTGTGAGATACCAAGTCTACTGCTATGGCGGTGGGAGT	1800
Qy	1801	GGCATTTGCCAACCTTTACAGACCTTATCAAGCTCAAGTGGTCTGTGCAAGTATTTATCA	1860
Db	1801	GGCATTTGCCAACCTTTACAGACCTTATCAAGCTCAAGTGGTCTGTGCAAGTATTTATCA	1860
Qy	1861	CTGAGACTCCGAGTCCAGCCCAACTCCACCCCATCCAGTGGAAATGCCACGAGCCATCTC	1920
Db	1861	CTGAGACTCCGAGTCCAGCCCAACTCCACCCCATCCAGTGGAAATGCCACGAGCCATCTC	1920
Qy	1921	ACATTTCCCAAGTACATTTCTCAGGTGGAGACCTGTGAGTATCCCAACCCAGAAACCTTGGAT	1980
Db	1921	ACATTTCCCAAGTACATTTCTCAGGTGGAGACCTGTGAGTATCCCAACCCAGAAACCTTGGAT	1980
Qy	1981	ACTGAGTCTCCTAATCTTTATCAATTTCTGATGGTTCCTTTTTTCCAGCTTTTGAGCCAA	2040
Db	1981	ACTGAGTCTCCTAATCTTTATCAATTTCTGATGGTTCCTTTTTTCCAGCTTTTGAGCCAA	2040
Qy	2041	CAACTCTGATTAACCTATTTCTTATAGCATTTTACTATATTTTGTAGTGAACAAATATG	2100
Db	2041	CAACTCTGATTAACCTATTTCTTATAGCATTTTACTATATTTTGTAGTGAACAAATATG	2100

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2005, 12:35:07 ; Search time 48 Seconds
(without alignments)
1105.787 Million cell updates/sec

Title: US-09-581-651D-2
Perfect score: 3687
Sequence: 1 MLRGPGLLLAVQCLGTA.....ISKYLRRPVSIPRNLGY 642

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3595.5	97.5	2386	2	US-09-961-403-1
2	3579.5	97.1	2386	1	US-09-016-366A-12
3	3578.5	97.1	2446	1	US-08-551-356-2
4	3578.5	97.1	2446	4	PCT-US93-12687-2
5	3439.5	93.3	2355	2	US-10-360-101-235
6	3433.5	93.1	2327	6	5455158-1
7	3422.5	92.8	2324	1	US-08-283-857-1
8	3422.5	92.8	2324	4	PCT-US95-09819-1
9	3406.5	92.4	2231	1	US-08-153-799-16
10	371	10.1	65	1	US-08-982-597A-19
11	371	10.1	65	2	US-09-136-218-19
12	298	8.1	663	2	US-09-194-468A-30
13	297	8.1	631	2	US-08-448-489-17
14	297	8.1	631	2	US-09-689-730-17
15	297	8.1	660	2	US-08-704-711A-18
16	297	8.1	660	2	US-09-521-220-18
17	297	8.1	660	2	US-09-391-104-19
18	297	8.1	660	2	US-09-917-254-89
19	297	8.1	660	2	US-09-949-016-6512
20	297	8.1	660	2	US-09-949-016-7937
21	297	8.1	660	2	US-10-153-185-14
22	295	8.0	422	1	US-08-836-854-12
23	294.5	8.0	188	1	US-08-142-449B-14
24	294.5	8.0	474	1	US-08-836-854-9
25	292.5	7.9	429	2	US-09-194-468A-45
26	290	7.9	48	1	US-08-982-597A-22
27	290	7.9	48	2	US-09-136-218-22

28	285.5	7.7	707	2	US-08-704-711A-19	Sequence 19, Appl
29	285.5	7.7	707	2	US-09-521-220-19	Sequence 19, Appl
30	285.5	7.7	707	2	US-09-391-104-20	Sequence 20, Appl
31	285.5	7.7	707	2	US-09-949-016-6575	Sequence 6575, Ap
32	285.5	7.7	708	2	US-08-448-489-16	Sequence 16, Appl
33	285.5	7.7	708	2	US-09-689-730-16	Sequence 16, Appl
34	285.5	7.7	713	2	US-09-949-016-10629	Sequence 10629, A
35	278	7.5	48	1	US-08-982-597A-20	Sequence 20, Appl
36	278	7.5	48	2	US-09-136-218-20	Sequence 20, Appl
37	258	7.0	43	1	US-08-982-597A-21	Sequence 21, Appl
38	258	7.0	43	2	US-09-136-218-21	Sequence 21, Appl
39	220	6.0	96	1	US-08-717-169-2	Sequence 2, Appl
40	220	6.0	96	2	US-09-228-901A-2	Sequence 2, Appl
41	219.5	6.0	60	1	US-08-982-597A-18	Sequence 18, Appl
42	219.5	6.0	60	2	US-09-136-218-18	Sequence 18, Appl
43	208	5.6	1455	2	US-08-840-062-5	Sequence 5, Appl
44	206	5.6	37	1	US-08-982-597A-17	Sequence 17, Appl
45	206	5.6	37	2	US-09-136-218-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-961-403-1
; Sequence 1, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961.403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-1

Query Match	97.5%	Score 3595.5	DB 2	Length 2386
Best Local Similarity	97.2%	Pred. No. 2.6e-301	Mismatches 2	Indels 15
Matches 629	Conservative 1			Gaps 1
Qy	1	MLRGPGLLLAVQCLGTA	VPSTGASKSRQAQMVQPSQSPVAVSQSKPCYDNGKHQ	60
Db	1	MLRGPGLLLAVQCLGTA	VPSTGASKSRQAQMVQPSQSPVAVSQSKPCYDNGKHQ	60
Qy	61	INQWERTYLGALVCTCYGSGRGNFCESKPEABETCFDKYTGNTYRVGDTYERPKDSMI	120	
Db	61	INQWERTYLGALVCTCYGSGRGNFCESKPEABETCFDKYTGNTYRVGDTYERPKDSMI	120	
Qy	121	WDCTCIGAGRGISCTIANRCHEGGQSYKIGDTRWRPHETGGYMLCVCCLGNGKGEWTK	180	
Db	121	WDCTCIGAGRGISCTIANRCHEGGQSYKIGDTRWRPHETGGYMLCVCCLGNGKGEWTK	180	
Qy	181	PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGSGGRITCTSRNRCNDQDTRTSY	240	
Db	181	PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGSGGRITCTSRNRCNDQDTRTSY	240	
Qy	241	RIGDTWSKONRGNLLQCICTGNRGKWKCEHRTSVQTTSSGSGPFTDVAAYVQOPHP	300	
Db	241	RIGDTWSKONRGNLLQCICTGNRGKWKCEHRTSVQTTSSGSGPFTDVAAYVQOPHP	300	
Qy	301	QPPYGHCVTDSGVVYSGWQLTKQNKQLCTCLNGVSCORTAVTQTYGNGNSGEP	360	

Db 301 PPPYGHCVTDGSGVYVSGMWLKTQGNKQMLCTCLGNGVSCQETAVTQTYGNSNGEPC 360
Qy 361 VLPFTYNDRT-----DSTTSNYEQDKYSFCTDHTVLVOTRGNSGALC 405
Db 361 VLPFTYNGRTFYSCCTTEGRQDGHLMWCTTSNYEQDKYSFCTDHTVLVOTRGNSGALC 420
Qy 406 HFPFLYNNHNTYDCTSEGRDNMKWCGTTQNYDADQKFGFCFMAAHEEICTTNEGVMYRI 465
Db 421 HFPFLYNNHNTYDCTSEGRDNMKWCGTTQNYDADQKFGFCFMAAHEEICTTNEGVMYRI 480
Qy 466 GDQWDKQDHMGHMRCTCVNGRGSWTCTIAYSQLRDQCIYVDDITVYNNVDTFKRHEEGHM 525
Db 481 GDQWDKQDHMGHMRCTCVNGRGSWTCTIAYSQLRDQCIYVDDITVYNNVDTFKRHEEGHM 540
Qy 526 LNCTCFGQGRGWKCDPVDQODSTGTFFYQIGDSWEKYVHGVRYQCYCYGRGIGEWHCQ 585
Db 541 LNCTCFGQGRGWKCDPVDQODSTGTFFYQIGDSWEKYVHGVRYQCYCYGRGIGEWHCQ 600
Qy 586 PLQTPSSSGPVEVITETPSQPNSHPIQWNAPOPSHISKYLWRP 632
Db 601 PLQTPSSSGPVEVITETPSQPNSHPIQWNAPOPSHISKYLWRP 647

RESULT 2

US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2386 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-016-366A-12

Query Match 97.1%; Score 3579.5; DB 1; Length 2386;
Best Local Similarity 96.8%; Pred. No. 6.3e-300;
Matches 626; Conservative 1; Mismatches 5; Indels 15; Gaps 1;

Qy 1 MLRGGPGLLLAVQCLGTAVPSTGASKSKRQAMQVQPOS PVAVSQSKPGCYDNGKHQY 60
Db 1 MLRGGPGLLLAVLCLGTAVPSTGASKSKRQAMQVQPOS PVAVSQSKPGCYDNGKHQY 60
Qy 61 INQWERTYLGNAUVCTCYGSGRGNFNCESPEABETCFDKYTGNTYRVGDTYERPKDSMI 120
Db 61 INQWERTYLGNAUVCTCYGSGRGNFNCESPEABETCFDKYTGNTYRVGDTYERPKDSMI 120
Qy 121 WDCTCICAGGRISCTTANRCHEGGSYKIGDTRRRPHETGGYMLECVCLGNGKGEWTCK 180
Db 121 WDCTCICAGGRISCTTANRCHEGGSYKIGDTRRRPHETGGYMLECVCLGNGKGEWTCK 180
Qy 181 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
Db 181 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
Qy 241 RIGDTWSKKNRGNLLQICITGNGRGWKERHTSVQTTSSGSGPFTDRAAAYQPOPH 300
Db 241 RIGDTWSKKNRGNLLQICITGNGRGWKERHTSVQTTSSGSGPFTDRAAAYQPOPH 300
Qy 301 PPPYGHCVTDGSGVYVSGMWLKTQGNKQMLCTCLGNGVSCQETAVTQTYGNSNGEPC 360
Db 301 PPPYGHCVTDGSGVYVSGMWLKTQGNKQMLCTCLGNGVSCQETAVTQTYGNSNGEPC 360
Qy 361 VLPFTYNDRT-----DSTTSNYEQDKYSFCTDHTVLVOTRGNSGALC 405
Db 361 VLPFTYNGRTFYSCCTTEGRQDGHLMWCTTSNYEQDKYSFCTDHTVLVOTRGNSGALC 420
Qy 406 HFPFLYNNHNTYDCTSEGRDNMKWCGTTQNYDADQKFGFCFMAAHEEICTTNEGVMYRI 465
Db 421 HFPFLYNNHNTYDCTSEGRDNMKWCGTTQNYDADQKFGFCFMAAHEEICTTNEGVMYRI 480
Qy 466 GDQWDKQDHMGHMRCTCVNGRGSWTCTIAYSQLRDQCIYVDDITVYNNVDTFKRHEEGHM 525
Db 481 GDQWDKQDHMGHMRCTCVNGRGSWTCTIAYSQLRDQCIYVDDITVYNNVDTFKRHEEGHM 540
Qy 526 LNCTCFGQGRGWKCDPVDQODSTGTFFYQIGDSWEKYVHGVRYQCYCYGRGIGEWHCQ 585
Db 541 LNCTCFGQGRGWKCDPVDQODSTGTFFYQIGDSWEKYVHGVRYQCYCYGRGIGEWHCQ 600
Qy 586 PLQTPSSSGPVEVITETPSQPNSHPIQWNAPOPSHISKYLWRP 632
Db 601 PLQTPSSSGPVEVITETPSQPNSHPIQWNAPOPSHISKYLWRP 647

RESULT 3
US-08-551-356-2
; Sequence 2, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2005, 18:06:49 ; Search time 405 Seconds
(without alignments)
9423.271 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaactggggcaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/1_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/H_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/PCUS_COMB.seq.*
- 7: /cgn2_6/ptodata/1/ina/PP_COMB.seq.*
- 8: /cgn2_6/ptodata/1/ina/RE_COMB.seq.*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1887.4	87.9	8044	3	US-09-566-921-135
2	1826.8	85.1	7803	2	US-08-551-356-1
3	1826.8	85.1	7803	6	PCT-US93-12687-1
4	1743.4	81.2	7679	3	US-09-220-132-38
5	1743.4	81.2	7680	3	US-09-023-655-1289
6	1743.4	81.2	7680	6	PCT-US93-09819-6
7	1740.2	81.1	7705	2	US-08-259-569-16
8	1740.2	81.1	7705	2	US-08-826-885-16
9	1738.6	81.0	7705	9	5455158-2
10	671.2	31.3	986	2	US-07-637-250A-8
11	671.2	31.3	986	2	US-08-145-061-8
12	97.4	4.5	186	2	US-08-153-799-5
13	73.8	3.4	2109	3	US-09-799-451-345
14	73.8	3.4	2334	3	US-09-023-655-996
15	73.8	3.4	2334	3	US-09-949-016-704
16	73.8	3.4	2335	3	US-09-799-451-346
17	73.8	3.4	2335	3	US-09-949-016-4758
18	73.2	3.4	11665	3	US-09-949-016-12446
19	73.2	3.4	11665	3	US-09-949-016-16500
20	72.4	3.4	85	2	US-08-259-569-28
21	72.4	3.4	85	2	US-08-826-885-28
22	72.2	3.4	2334	2	US-08-457-304A-33
23	72.2	3.4	2334	2	US-08-456-701A-33
24	72.2	3.4	2334	3	US-08-684-932A-33

25	72	3.4	72	2	US-08-259-569-29	Sequence 29, Appl
26	72	3.4	72	2	US-08-826-885-29	Sequence 29, Appl
27	72	3.4	72	9	5455158-9	Patent No. 5455158
28	71.8	3.3	1983	3	US-09-949-016-2066	Sequence 2066, Ap
29	71.8	3.3	3069	3	US-09-949-016-641	Sequence 641, App
30	69	3.2	69	2	US-08-259-569-23	Sequence 23, Appl
31	69	3.2	69	2	US-08-826-885-23	Sequence 23, Appl
32	65.8	3.1	69	9	5455158-6	Patent No. 5455158
33	65.4	3.0	69	2	US-08-259-569-30	Sequence 30, Appl
34	65.4	3.0	69	2	US-08-826-885-30	Sequence 30, Appl
35	65.2	3.0	601	3	US-09-949-016-30391	Sequence 30391, A
36	65.2	3.0	601	3	US-09-949-016-70165	Sequence 70165, A
37	65.2	3.0	29954	3	US-09-949-016-13808	Sequence 13808, A
38	65.2	3.0	31040	3	US-09-949-016-12383	Sequence 12383, A
39	65	3.0	75	2	US-08-259-569-18	Sequence 18, Appl
40	65	3.0	75	2	US-08-826-885-18	Sequence 18, Appl
41	64.8	3.0	68	9	5455158-4	Patent No. 5455158
42	61.8	2.9	71	2	US-08-259-569-27	Sequence 27, Appl
43	61.8	2.9	71	2	US-08-826-885-27	Sequence 27, Appl
44	61.8	2.9	71	9	5455158-8	Patent No. 5455158
45	61.2	2.9	71	2	US-08-259-569-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-566-921-135
; Sequence 135, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne P.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 135
; LENGTH: 8044
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 427813.14
US-09-566-921-135

Query Match	87.9%	Score	1887.4	DB	3	Length	8044
Best Local Similarity	97.4%	Pred. No.	0	Mismatches	0	Indels	45
Matches	1946	Conservative	0	Gaps	1		
Oy	1	CAAACTTGGTGGCAACTTGGCTCCCGTGGGGGGTCTCTCCCGCCAGCGTCTCAACATGC	60				
Db	213	CAAACTTGGTGGCAACTTGGCTCCCGTGGGGGGTCTCTCCCGCCAGCGTCTCAACATGC	272				
Oy	61	TTAGGGGTCCGGGGCCCGGGTCTGCTGCTGCTGGCCGCTCCAGTGTCTGGGACACGGGTGC	120				
Db	273	TTAGGGGTCCGGGGCCCGGGTCTGCTGCTGCTGGCCGCTCTGCTGCTGGGACACGGGTGC	332				
Oy	121	CTCTCCAGGAGGCTCGAAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTTCAGCCCCAGTCCC	180				
Db	333	CTCTCCAGGAGGCTCGAAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTTCAGCCCCAGTCCC	392				
Oy	181	CGTGGTGTGCTCAGTCAAGCAAGCCCGTTGTTATGCAATGGAACACATATCAGATAA	240				
Db	393	CGTGGTGTGCTCAGTCAAGCAAGCCCGTTGTTATGCAATGGAACACATATCAGATAA	452				
Oy	241	ATCAACAGTGGGAGCGGACCTACTAGGCAATGGTTGGTTGTACTTGTATGGAGAA	300				
Db	453	ATCAACAGTGGGAGCGGACCTACTAGGCAATGGTTGGTTGTACTTGTATGGAGAA	512				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:49:16 ; Search time 664.314 Seconds
(without alignments)
1625.652 Million cell updates/sec

Title: US-09-581-651D-1
Perfect score: 3781
Sequence: 1 NLVATCLPVRSLPHRLNML.....ISKYILRWPRVSPRNLGY 660

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool_h/US09581651/runat_07112005_092224_28801/app.query.fasta_1.1834
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORES=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651 -CGN 1 1 187 -runat_07112005_092224_28801 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3686.5	97.5	8044	4	US-09-566-921-135
2	3584.5	94.8	7803	2	Sequence 135, App
3	3584.5	94.8	7803	5	Sequence 1, Appli
4	3448.5	91.2	7679	4	Sequence 38, Appl
5	3448.5	91.2	7680	4	Sequence 1289, Ap
6	3448.5	91.2	7680	5	Sequence 6, Appli
7	3433.5	90.8	7705	1	Sequence 16, Appl
8	3433.5	90.8	7705	2	Sequence 16, Appl
9	3428.5	90.7	7705	6	Patent No. 5455158
10	3428.5	90.7	7705	6	Patent No. 5455158
11	1665.5	44.0	986	1	US-07-637-250A-8
12	1665.5	44.0	986	1	US-07-637-250A-8

13	325.5	8.6	2109	4	US-09-799-451-345	Sequence 345, App
14	311	8.2	2335	1	US-09-799-451-346	Sequence 346, App
15	306.5	8.1	2333	1	US-08-392-678-33	Sequence 33, Appl
16	305.5	8.1	2334	1	US-08-457-304A-33	Sequence 33, Appl
17	305.5	8.1	2334	1	US-08-456-701A-33	Sequence 33, Appl
18	305.5	8.1	2334	3	US-08-684-932A-33	Sequence 33, Appl
19	305.5	8.1	2334	4	US-09-023-655-996	Sequence 996, App
20	305.5	8.1	2334	4	US-09-949-016-704	Sequence 704, App
21	305.5	8.1	2335	4	US-09-949-016-4758	Sequence 4758, Ap
22	302	8.0	1983	4	US-09-949-016-2066	Sequence 2066, Ap
23	302	8.0	3069	4	US-09-949-016-641	Sequence 641, App
24	298	7.9	2123	4	US-08-194-468A-29	Sequence 29, Appl
25	294.5	7.8	567	1	US-08-142-449B-5	Sequence 5, Appli
26	225	6.0	11665	4	US-09-949-016-12446	Sequence 12446, A
27	225	6.0	11665	4	US-09-949-016-16500	Sequence 16500, A
28	217.5	5.8	4588	3	US-08-840-062-1	Sequence 1, Appli
29	208	5.5	4627	4	US-09-949-016-935	Sequence 935, App
30	208	5.5	5633	4	US-09-023-655-1490	Sequence 1490, Ap
31	199.5	5.3	4852	4	US-09-976-594-167	Sequence 167, App
32	199.5	5.3	5158	4	US-09-023-655-1347	Sequence 1347, Ap
33	197	5.2	4621	4	US-09-949-016-4577	Sequence 4577, Ap
34	197	5.2	5627	4	US-09-949-016-2444	Sequence 2444, Ap
35	196.5	5.2	29954	4	US-09-949-016-13808	Sequence 13808, A
36	196.5	5.2	31040	4	US-09-949-016-12383	Sequence 12383, A
37	193.5	5.1	4464	2	US-08-400-159-7	Sequence 7, Appli
38	193.5	5.1	4483	3	US-08-611-729A-7	Sequence 7, Appli
39	193.5	5.1	4483	4	US-09-195-524-7	Sequence 7, Appli
40	193	5.1	3955	3	US-09-214-278-4	Sequence 4, Appli
41	193	5.1	3955	4	US-09-855-722-4	Sequence 4, Appli
42	189.5	5.0	4771	3	US-08-840-062-3	Sequence 3, Appli
43	189	5.0	601	4	US-09-949-016-30388	Sequence 30388, A
44	189	5.0	601	4	US-09-949-016-30391	Sequence 30391, A
45	189	5.0	601	4	US-09-949-016-70162	Sequence 70162, A

ALIGNMENTS

RESULT 1

US-09-566-921-135
Sequence 135, Application US/09566921
Patent No. 6682888
GENERAL INFORMATION:
APPLICANT: Loring, Jeanne F.
APPLICANT: Tingley, Debora W.
APPLICANT: Edwards, Carla M.
TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
FILE REFERENCE: PA-0024 US
CURRENT APPLICATION NUMBER: US/09/566, 921
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PERL Program
SEQ ID NO 135
LENGTH: 8044
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6682888 427813.14
US-09-566-921-135

Alignment Scores:
Pred. No.: 0
Score: 3686.50
Percent Similarity: 97.29%
Best Local Similarity: 97.29%
Query Match: 97.50%
DB: 4
Matches: 8044
Conservative: 647
Mismatches: 3
Indels: 15
Gaps: 1

US-09-581-651D-1 (1-660) x US-09-566-921-135 (1-8044)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAenMetLeu 20

Db 215 AACTTGTGTGCACTTGCCTCCCGGCGGGGTCTCTCCCCACCCTCTCAACATGCTT 274

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2005, 18:42:04 ; Search time 1153 Seconds
(without alignments)
11023.144 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaacttggtggcaactgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2147	100.0	2147	2	AX81299 Human mig
2	2143.8	99.9	2358	13	ADR67201 Human bla
3	2119.8	98.7	2443	13	Adq38575 Human SNP
4	2119	98.7	2127	10	Adi18477 Human pro
5	2064.8	96.2	2488	13	Adq38581 Human SNP
6	1975.8	92.0	4295	8	ACC72037 BCU0770 g
7	1925.8	89.7	1929	13	Adsi17488 Nucleotid
8	1925.8	89.7	1929	13	Adr97657 Human fib
9	1890.6	88.1	7242	13	Adr66637 Human pro
10	1890.6	88.1	7242	13	Adr65953 Human pro
11	1890.6	88.1	7550	8	ACC00412 Human cel
12	1890.6	88.1	8027	11	Adp64998 Human fib
13	1890.6	88.1	8027	12	ADG89565 Human fib
14	1890.6	88.1	8027	13	Adq29601 Human col
15	1890.6	88.1	8027	13	Adr67200 Human bla
16	1889.8	88.0	6510	13	Adq38578 Human SNP
17	1889.8	88.0	7823	13	Adq38582 Human SNP
18	1889.8	88.0	7848	13	Adq38583 Human SNP
19	1889.8	88.0	7935	13	Adq38577 Human SNP
20	1889.8	88.0	7959	13	Adq38586 Human SNP

21	1889.8	88.0	8013	13	ADQ38579	Adq38579 Human SNP
22	1889.8	88.0	8155	13	ADQ38584	Adq38584 Human SNP
23	1889.8	88.0	8226	13	ADQ38574	Adq38574 Human SNP
24	1889.8	88.0	8278	13	ADQ38587	Adq38587 Human SNP
25	1889.8	88.0	8332	13	ADQ38576	Adq38576 Human SNP
26	1889.8	88.0	8371	13	ADQ38580	Adq38580 Human SNP
27	1887.4	87.9	7795	10	ADJ56196	Adj56196 Zebrafish
28	1887.4	87.9	7867	4	AAI57803	AAI57803 Human pol
29	1887.4	87.9	8044	6	AAI57803	AAI57803 Human pol
30	1887.4	87.9	8044	12	ADI61767	ADI61767 Human DNA
31	1887.4	87.9	8062	4	AAI57802	AAI57802 Human pol
32	1887.4	87.9	8137	4	AAI57805	AAI57805 Human pol
33	1887.4	87.9	8230	4	AAI57804	AAI57804 Human pol
34	1858.6	86.6	8066	4	AAI59591	AAI59591 Human pol
35	1858.6	86.6	8066	4	AAI59589	AAI59589 Human pol
36	1858.6	86.6	8066	4	AAI59588	AAI59588 Human pol
37	1858.6	86.6	8066	4	AAI59590	AAI59590 Human pol
38	1839.6	85.7	7361	9	ACD06170	ACD06170 Human GDN
39	1838.8	85.6	8039	3	AAI59889	AAI59889 Human FN
40	1826.8	85.1	7803	2	AAQ70009	AAQ70009 Fibrinoge
41	1814.6	84.5	8216	5	AAI59591	AAI59591 Human pol
42	1754.6	81.7	7677	13	ACN37820	ACN37820 Tumour-as
43	1743.4	81.2	4860	3	AAA35009	AAA35009 Human ade
44	1743.4	81.2	6988	9	ACD06169	ACD06169 Human GDN
45	1743.4	81.2	7020	13	ADR90518	ADR90518 Human GDN

ALIGNMENTS

RESULT 1
AX81299
ID AX81299 standard; cdna; 2147 BP.

XX AX81299;

DT 21-SEP-1999 (first entry)

DE Human migration stimulating factor (MSF) 1-alpha encoding cdna.

KW Migration stimulatory factor; MSF; cell migration; modulation; human;
wound healing; scarring; MSF1-alpha; 88.

OS Homo sapiens.

PN WO9931233-A1.

XX 24-JUN-1999.

PF 15-DEC-1998; 98WO-GB003766.

PR 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

DR WPI; 1999-430039/36.

DR P-PSDB; RAY28901.

XX Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

PS Example 1; Fig 1; 86pp; English.

CC The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. The present sequence CC represents the nucleotide sequence encoding a human MSF1-alpha protein

XX Sequence 2147 BP; 575 A; 512 C; 575 G; 485 T; 0 U; 0 Other;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2005, 20:50:35 ; Search time 9269 Seconds
(without alignments)
11223.794 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaacttggcgcaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY NUC

Gapop 10'0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2147	100.0	2147	6	BD137021
2	2147	100.0	2147	6	AX003229
3	2147	100.0	2147	9	HSA276395
4	2143.8	99.9	2358	6	C0875358
5	2084	97.1	2192	9	HSA535086
6	2063.4	96.1	4321	9	HSM806214
7	1925.8	89.7	1929	6	C0871810
8	1925.8	89.7	1929	6	C0871828
9	1890.6	88.1	7753	9	AB191261
10	1890.6	88.1	8027	6	C0833991
11	1890.6	88.1	8027	6	C0875357
12	1889	88.0	7868	9	HSM806653
13	1889	88.0	7951	9	HSM806901
14	1889	88.0	7951	9	HSM806902
15	1889	88.0	8411	9	HSM806992
16	1887.4	87.9	7502	9	HSM804082
17	1887.4	87.9	8042	9	HSM806805
18	1887.4	87.9	8044	6	AK454662
19	1887.4	87.9	8044	6	AX281712

20	1885.8	87.8	7544	9	HSM803509
21	1885.8	87.8	8320	9	HSM806267
22	1884.2	87.8	8030	9	HSM806170
23	1884.2	87.8	8035	9	HSM806171
24	1884.2	87.8	8421	9	HSM806903
25	1826.8	85.1	7803	6	AR051657
26	1743.4	81.2	7679	6	AR274901
27	1743.4	81.2	7680	6	AR380744
28	1743.4	81.2	7680	6	AX277596
29	1743.4	81.2	7680	6	AX335368
30	1743.4	81.2	7680	9	HSFIB1
31	1740.2	81.1	7705	6	AL14133
32	1740.2	81.1	7705	6	AR034630
33	1740.2	81.1	7705	6	E011162
34	1740.2	81.1	7705	6	E011162
35	1738.6	81.0	7705	6	AR364992
36	1547	72.1	6816	6	CQ715726
37	1461.8	68.1	3059	10	BC051082
38	1429.8	66.6	8329	6	AX402055
39	1429.8	66.6	8329	10	RNFIBRON
40	1263	58.8	7323	11	BV178397
41	1027.8	47.9	1139	9	HSU42404
42	990.6	46.1	2409	6	CQ731571
43	975.4	45.4	8313	5	BC072841
44	972.2	45.3	8216	5	XELFBRNCT
45	940	43.8	7299	9	HSM807162

ALIGNMENTS

RESULT 1
BD137021
LOCUS BD137021 2147 bp DNA linear PAT 18-SEP-2002
DEFINITION Polypeptides, polynucleotides and uses thereof.
ACCESSION BD137021
VERSION BD137021.1 GI:23231966
KEYWORDS JP 2002508179-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2147)
AUTHORS Schor,S.L. and Schor,A.M.
TITLE Polypeptides, polynucleotides and uses thereof
JOURNAL Patent: JP 2002508179-A 1 19-MAR-2002;
UNIVERSITY OF DUNDEE
COMMENT OS Homo sapiens (human)
PN JP 2002508179-A/1
PD 19-MAR-2002
PF 15-DEC-1997 GB 9726539.1
PR 16-DEC-1997 GB 9726539.1
PI SETH LAWRENCE SCHOR,ANA MARIA SCHOR
PC C12N15/09,A61K39/00,A61F17/02,C07K14/78,C07K16/18,C12N5/10, PC
C12P21/08,
PC C12Q1/68,G01N33/574,C12N15/00,A61K37/02,C12N5/00 CC
Polypeptides, polynucleotides and uses thereof. FH Key
Location/Qualifiers
FT source 1..2147
FT Location/Qualifiers
1..2147
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES

source
1..2147
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2147; DB 6; Length 2147;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CAAACTTGGTGGCACTTGGCTCCCGTGGGGGCTCTCCCGCCACGCTCAACATGC 60
|||||

Db 1 CAAACTTGGTGGCAACTTGGCTCCCGGTGGGGCGTCTCTCCCCCAACGGTCTCAACATGC 60
Qy
61 TTAGGGGTCCGGGGCCCGGGCTGTCTGTCTGGCGGTCCAGTGGCTGGGGACAGCGGTGC 120
Db
61 TTAGGGGTCCGGGGCCCGGGCTGTCTGTCTGGCGGTCCAGTGGCTGGGGACAGCGGTGC 120
Qy
121 CCTCCACGGGAGCTCGAAGAGCAAGAGGCAAGGCTCAGCAAAATGTTTCAGGCCCAAGTCCC 180
Db
121 CCTCCACGGGAGCTCGAAGAGCAAGAGGCAAGGCTCAGCAAAATGTTTCAGGCCCAAGTCCC 180
Qy
181 CGGTGGTGTGAGTCAAGCAAGCCGGTGTATGCAAAATGTTTCAGGCCCAAGTCCC 240
Db
181 CGGTGGTGTGAGTCAAGCAAGCCGGTGTATGCAAAATGTTTCAGGCCCAAGTCCC 240
Qy
241 ATCAACAGTGGGAGCGGACCTACCTAGGCAATGGTGTGTTTGTACTTGTATGGAGGAA 300
Db
241 ATCAACAGTGGGAGCGGACCTACCTAGGCAATGGTGTGTTTGTACTTGTATGGAGGAA 300
Qy
301 GCCGAGGTTTTAACTGCCGAGAGTAACTGAAGCTGAAGAGACTTGGCTTTGACCAAGTACA 360
Db
301 GCCGAGGTTTTAACTGCCGAGAGTAACTGAAGCTGAAGAGACTTGGCTTTGACCAAGTACA 360
Qy
361 CTGGGAAACATTAACCGAGTGGGTGACACTTATGAGCGTCTTAAGAGACTCCATGATCTGGG 420
Db
361 CTGGGAAACATTAACCGAGTGGGTGACACTTATGAGCGTCTTAAGAGACTCCATGATCTGGG 420
Qy
421 ACTGTACCTGATCCGGGCTGGGCGAGGAGAAATGAGCTGTACCTGCAAAACCGCTGCC 480
Db
421 ACTGTACCTGATCCGGGCTGGGCGAGGAGAAATGAGCTGTACCTGCAAAACCGCTGCC 480
Qy
481 ATGAAGGGGTGAGTCTTACCAAGATTGGTGACACCTGGAGGAGACCAATGAGACTGTGTG 540
Db
481 ATGAAGGGGTGAGTCTTACCAAGATTGGTGACACCTGGAGGAGACCAATGAGACTGTGTG 540
Qy
541 GTTACATGTTAGAGTGTGTGTCTTGTGTAATGGAAGAGAGAAATGGAACCTGCAAGCCCA 600
Db
541 GTTACATGTTAGAGTGTGTGTCTTGTGTAATGGAAGAGAGAAATGGAACCTGCAAGCCCA 600
Qy
601 TAGCTGAGAAGTGTGATCATGCTGTGGGACTTCTCTATGTGTCTGGAGAAAGCTGGG 660
Db
601 TAGCTGAGAAGTGTGATCATGCTGTGGGACTTCTCTATGTGTCTGGAGAAAGCTGGG 660
Qy
661 AGAAGCCCTCAACAGGCTGGATGATGATGATGTTGCTGGGAGAGGACGACGAC 720
Db
661 AGAAGCCCTCAACAGGCTGGATGATGATGATGTTGCTGGGAGAGGACGACGAC 720
Qy
721 GCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTCTATGAA 780
Db
721 GCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTCTATGAA 780
Qy
781 TTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGTCTCAGTGCATCTGCACAG 840
Db
781 TTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGTCTCAGTGCATCTGCACAG 840
Qy
841 GCAACGGCCGAGGAGTGGAAAGTGTGAGAGGACACCTCTGTGAGACCAACATCGAGCG 900
Db
841 GCAACGGCCGAGGAGTGGAAAGTGTGAGAGGACACCTCTGTGAGACCAACATCGAGCG 900
Qy
901 GATCTGGCCCTTACCGGATGTTCTGTGACAGTGTGTTTACCAACCGGAGGCTCACCCCGAGC 960
Db
901 GATCTGGCCCTTACCGGATGTTCTGTGACAGTGTGTTTACCAACCGGAGGCTCACCCCGAGC 960
Qy
961 CTCCTCCCTATGGCCACTGTGTACAGAGTGGTGTGTCTACTCTGTGGGGATGCAGT 1020
Db
961 CTCCTCCCTATGGCCACTGTGTACAGAGTGGTGTGTCTACTCTGTGGGGATGCAGT 1020
Qy
1021 GGCTGAAGACACAGGAAATAGCAAAATGCTTTTGACAGTGGCTGGGCAACGAGTCACT 1080
Db
1021 GGCTGAAGACACAGGAAATAGCAAAATGCTTTTGACAGTGGCTGGGCAACGAGTCACT 1080
Qy
1081 GCCAAGAGACAGCTGTGTAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTGCT 1140
Db
1081 GCCAAGAGACAGCTGTGTAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTGCT 1140

Qy 1141 TACCATTACCTACACGACAGGACGACGACCAACTTCGAATTTATGAGCAGGACCAGA 1200
Db
1141 TACCATTACCTACACGACAGGACGACGACCAACTTCGAATTTATGAGCAGGACCAGA 1200
Qy 1201 AATACTCTTTCTGCACAGACCAACATGTGTTTGGTTTCAGACTCGAGGAGGAAAATTCCAATG 1260
Db
1201 AATACTCTTTCTGCACAGACCAACATGTGTTTGGTTTCAGACTCGAGGAGGAAAATTCCAATG 1260
Qy 1261 GTGCTTGTGCGCACTTCCCTTCTATACAAACCAACCAATTTACACTGATTTGCACTTCTG 1320
Db
1261 GTGCTTGTGCGCACTTCCCTTCTATACAAACCAACCAATTTACACTGATTTGCACTTCTG 1320
Qy 1321 AGGCGAAGAGACCAACATGAAGTGGTGGGACCAACAGAACTATGATGCGGACGACCA 1380
Db
1321 AGGCGAAGAGACCAACATGAAGTGGTGGGACCAACAGAACTATGATGCGGACGACCA 1380
Qy 1381 AGTTTGGGTTCTGCCCATGGCTGCCACGAGGAAATCTGCAAACTGAAGGGGTCA 1440
Db
1381 AGTTTGGGTTCTGCCCATGGCTGCCACGAGGAAATCTGCAAACTGAAGGGGTCA 1440
Qy 1441 TGTACCCGATTTGGAGATCAGTGGGATAGCAGCATGATGGGTCAATGATGAGGTGCA 1500
Db
1441 TGTACCCGATTTGGAGATCAGTGGGATAGCAGCATGATGGGTCAATGATGAGGTGCA 1500
Qy 1501 CGTGTGTTGGGAATGGTCTGGGGAATGGACATGCTATTCGCTACTCGCAGCTTCGAGATC 1560
Db
1501 CGTGTGTTGGGAATGGTCTGGGGAATGGACATGCTATTCGCTACTCGCAGCTTCGAGATC 1560
Qy 1561 AGTCGATTTGTTGATGACATCATTTCAATGTGAACGACATTCACCAAGCGTCATGAAG 1620
Db
1561 AGTCGATTTGTTGATGACATCATTTCAATGTGAACGACATTCACCAAGCGTCATGAAG 1620
Qy 1621 AGGGGCACTGCTGAACTGTACATGCTTCGGTCAAGGTCGGGGCAGGTGGAAGTGTGATC 1680
Db
1621 AGGGGCACTGCTGAACTGTACATGCTTCGGTCAAGGTCGGGGCAGGTGGAAGTGTGATC 1680
Qy 1681 CCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCAAATTTGGAGATTCATGGG 1740
Db
1681 CCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCAAATTTGGAGATTCATGGG 1740
Qy 1741 AGAAGTATGTCATGGTGTGACATGATGATGCTACTGCTATGGCGGTGGCATTTGGGAGT 1800
Db
1741 AGAAGTATGTCATGGTGTGACATGATGATGCTACTGCTATGGCGGTGGCATTTGGGAGT 1800
Qy 1801 GGCAATGCCAACCTTTACAGACCTTATCAAGCTCAAGTGGTCTGTCGAAGTATTTATCA 1860
Db
1801 GGCAATGCCAACCTTTACAGACCTTATCAAGCTCAAGTGGTCTGTCGAAGTATTTATCA 1860
Qy 1861 CTGAGACTCCGAGTCAGCCCAACTCCCAACCCATCCAGTGGAAATGCACCAAGCCATCTC 1920
Db
1861 CTGAGACTCCGAGTCAGCCCAACTCCCAACCCATCCAGTGGAAATGCACCAAGCCATCTC 1920
Qy 1921 ACATTTCCAAAGTACATTTCTCAGGTGGAGACTGTGAGTATCCCAACCAAGAACTTTGGAT 1980
Db
1921 ACATTTCCAAAGTACATTTCTCAGGTGGAGACTGTGAGTATCCCAACCAAGAACTTTGGAT 1980
Qy 1981 ACTGAGTCTCCTAATCTTATCAATTTGATGCTTTTCTTTTCCAGCTTTTGAGCCAA 2040
Db
1981 ACTGAGTCTCCTAATCTTATCAATTTGATGCTTTTCTTTTCCAGCTTTTGAGCCAA 2040
Qy 2041 CAATCTGATTAATCTTATCTATAGCATTTACTATATTTGTTTGTAGTGAACAAACATATG 2100
Db
2041 CAATCTGATTAATCTTATCTATAGCATTTACTATATTTGTTTGTAGTGAACAAACATATG 2100
Qy 2101 TGGTCAATTAATGACTTTGTAGACTGAAAAAATAAAAAAATAAAAAA 2147
Db
2101 TGGTCAATTAATGACTTTGTAGACTGAAAAAATAAAAAAATAAAAAA 2147

RESULT 2
AX003229
LOCUS

AX003229 2147 bp DNA linear PAT 24-AUG-2000

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 11, 2005, 03:57:22 ; Search time 199.5 Seconds

(without alignments)

11021.895 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 4096

Sequence: 1 caaacttggtgcacttgcc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO_spool_h/US09581651/runat_07112005_092113_27541/app_query.fasta_1.2311
-DB=UniProt -QWMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US09581651 @CGN 1.1 212 @runat_07112005_092113_27541 -NCPUS=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3732.5	91.1	749	2 Q7Z391	Q7z391 homo sapien
2	3686.5	90.0	1103	2 Q6MZ24	Q6mzf4 homo sapien
3	3685.5	90.0	2357	2 Q68D74	Q68dt4 homo sapien
4	3682.5	89.9	2444	2 Q6N025	Q6n025 homo sapien
5	3595.5	87.8	2386	1 F1NC HUMAN	P02751 homo sapien
6	3592.5	87.7	2296	2 Q6N0A6	Q6n0a6 homo sapien
7	3592.5	87.7	2477	2 Q6MZU5	Q6mzu5 homo sapien
8	3588.5	87.6	2267	2 Q68DP9	Q68dp9 homo sapien
9	3567.5	87.1	2240	2 Q68DP8	Q68dp8 homo sapien
10	3393.5	82.8	2477	1 F1NC MOUSE	P11276 mus musculus
11	3354.5	81.9	2477	1 F1NC RAT	P04937 rattus norv
12	3328.5	81.3	2265	1 F1NC BOVIN	P07589 bos taurus
13	2888	70.5	2481	2 Q6GQ25	Q6gqa5 xenopus lae
14	2878	70.3	2481	1 F1NC XENLA	Q91740 xenopus lae
15	2522	61.6	922	2 O93405	O93405 brachydanio
16	2522	61.6	2478	2 O93406	O93406 brachydanio

17	2330	56.9	2408	2 Q6JAN2	Q6jan2 brachydanio
18	1909.5	46.6	2193	2 Q6M2M7	Q6mzm7 homo sapien
19	1496	36.5	296	2 Q8C6J7	Q8c6j7 mus musculus
20	961	23.5	190	1 F1NC NOTVI	Q91400 notophthalm
21	750	18.3	141	2 Q90XQ2	Q90xq2 ambystoma m
22	543	13.3	215	2 Q6DD34	Q6dd34 xenopus lae
23	319	7.8	679	2 Q98856	Q98856 cynops pyrr
24	318	7.8	670	2 Q6DF16	Q6df16 xenopus tro
25	315.5	7.7	680	2 Q7T317	Q7t317 brachydanio
26	315	7.7	2146	2 Q68CX6	Q68cx6 homo sapien
27	313.5	7.7	673	2 Q90YB3	Q90yb3 paralicthy
28	310.5	7.6	674	2 Q98TC6	Q98tc6 cyprinus ca
29	310.5	7.6	690	2 Q9PVM5	Q9pvm5 oryzias lat
30	306	7.5	675	2 Q8QFQ6	Q8qfq6 oncorhynch
31	302.5	7.4	810	2 Q8R3F3	Q8r3f3 mus musculus
32	302	7.4	671	2 Q6PF33	Q6pf33 xenopus lae
33	302	7.4	671	2 Q9W7L6	Q9w7l6 xenopus lae
34	301	7.3	655	2 Q9W635	Q9w635 oncorhynch
35	301	7.3	661	2 Q95J44	Q95j44 sus scrofa
36	301	7.3	662	1 MM02 RABIT	P50757 oryctolagus
37	300.5	7.3	216	2 Q99KD0	Q99kd0 mus musculus
38	300.5	7.3	686	2 Q9DE15	Q9del5 gallus gall
39	300	7.3	662	1 MM02 MOUSE	P33434 mus musculus
40	300	7.3	662	1 MM02 RAT	P33436 rattus norv
41	300	7.3	707	1 MM09 RABIT	P41246 oryctolagus
42	300	7.3	708	1 MM09 RAT	P50282 rattus norv
43	298	7.3	412	2 Q71U44	Q71u44 bos taurus
44	298	7.3	663	1 MM02 CHICK	Q90611 gallus gall
45	297	7.3	660	1 MM02 HUMAN	P08253 homo sapien

ALIGNMENTS

RESULT 1

Q7Z391 PRELIMINARY; PRT; 749 AA.
ID Q7Z391
AC Q7Z391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp668B18150.
GN Name=DKFZp668B18150;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloeker H., Boescher M., Mewes H.W., Weil B., Amid C., Oanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -
DR HSSP; Q96KP7; IFR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 9.
DR PROSITE; PS00023; FIBRONECTIN 2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Alignment Scores:

Pred. No.: 6.05e-267 Length: 749

Score: 3732.50 Matches: 656

Percent similarity: 97.19% Conservative: 0

Best Local Similarity: 97.19%		Mismatches: 4
Query Match: 91.13%	Indels: 15	
DB: 2	Gaps: 1	
US-09-581-651D-3 (1-2147) x Q72391 (1-749)		
QY	3 AACTTGGTGGCAACTTGGCTCCCGGTGGGGCGTCTCTCCCGCCACCGCTCTCAACATGCTT	62
DB	75 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu	94
QY	63 AGGGTCCGGGCGCGGCTGCTGCTGCGCGTCCAGTGGCGGACAGCGGTGCC	122
DB	95 ArgGlyProGlyProGlyLeuLeuLeuAlaValLeuCysLeuGlyThrAlaValPro	114
QY	123 TCCAGGAGAGCTCGAAGAGCAGGAGGCTCAGCAATGGTTTCAGGCCCGCCCG	182
DB	115 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro	134
QY	183 GTGGCTGTCAAGCAAGCCCGTTGTTATGACAATGGAACACTATCAGATAAAT	242
DB	135 ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnLeu	154
QY	243 CAACAGTGGAGCGGACCTAGCTAGGCAATGCGTTGGTTGTATGGAGGAGC	302
DB	155 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlySer	174
QY	303 CGAGGTTTAACTGCGAGAGTAAACCTGACCTGAGCTGAAGACTTGTTCGACAGTACT	362
DB	175 ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysTyrThr	194
QY	363 GGGAACTTACCGAGTGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCGGGAC	422
DB	195 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp	214
QY	423 TGTACTGCATCGGGGCTGGCGAGGAGAAATAGCTGTACATCGCAACCGCTGCCAT	482
DB	215 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis	234
QY	483 GAAGGGGCTCAGTCTACAGATTGGTGACACCTGAGGAGGACCATGAGCTGTGGT	542
DB	235 GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgProHisIleGluThrGly	254
QY	543 TACATGTTAGAGTGTGTCTTGTGTAATGGAAGAGAGATGACCTCGAAGCCCAT	602
DB	255 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysProIle	274
QY	603 GCTGAGAGTGTGTTGATCATGCTGCTGGGACTTCTCTATGCTGGAGAAACGTGGAG	662
DB	275 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyLysThrTrpGlu	294
QY	663 AAGCCCTACCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	722
DB	295 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg	314
QY	723 ATCACTTGCACTTCTAGAAATAGATGACATGACATGACATGACATGACATGACATGAT	782
DB	315 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle	334
QY	783 GGAGACACTGGAGCAAGGATTAATCGAGAAACCTGCTCCAGTGCATGCTGCAAGGC	842
DB	335 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly	354
QY	843 AACGCCGAGGAGTGGAGTGGAGGACACCTCTGTCAGACACACATCGAGCGGA	902
DB	355 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly	374
QY	903 TCTGCGCCCTTACCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	962
DB	375 SerGlyProPheThrAspValArgAlaValTyrGlnProGlnProHisProGlnPro	394
QY	963 CTTCCCTATGGCCTATGCTGTCAGACAGTGTGTGTCTACTCTGTGGGGATGACAGTG	1022
DB	395 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp	414

RESULT 2

Q6MZF4

ID Q6MZF4

AC Q6MZF4;

PRELIMINARY;

PRT; 1103 AA.

QY	1023 CTGAAGACACAGGAAATAGCAAAATGCTTTGACGTGCTCTGGGCAACGGAGTCAGCTGC	1082
DB	415 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys	434
QY	1083 CAAGAGACAGCTGPAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTCTTA	1142
DB	435 GlnGluThrAlaValThrGlnThrTyrGlyLysAsnSerAsnGluGluProCysValLeu	454
QY	1143 CCATTACCTACACGACGACGACG-	1166
DB	455 ProPheThrTyrAsnGlyArgThrPheTyrSerCysThrThrGluGlyArgGlnAspGly	474
QY	1167 -----GACAGCAACAACCTTCCAAATTATGAGCAGGACGAGCAAAATCTCTTCTGCACA	1217
DB	475 HisLeuTrpCysSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr	494
QY	1218 GACCACACTGTTTGGTTCAGACTCGAGAGGAAATCCAAATGTGTGCTTGTGCCACTTC	1277
DB	495 AspHisThrValLeuValGlnThrArgGlyLysAsnSerAsnGlyAlaLeuCysHisPhe	514
QY	1278 CCCTTCTATACACCAACCAATTACACTGATTGCACTTCTGAGGGCAGAGAGACAAAC	1337
DB	515 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgArgAspAsn	534
QY	1338 ATGAAGTGGTGGGACCCACACAGAACTATGATGCCGACCAAGAGTGTGGTCTCTGCCCC	1397
DB	535 MetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro	554
QY	1398 ATGCTGCCACAGGAAATCTGCACAACTGAAGGGGTGATGATCCCGATGGAGAT	1457
DB	555 MetAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp	574
QY	1458 CAGTGGGATAGCAGCATGATGCGGTGCACATGATGAGTGCAGTGTGTTGGGAATGCT	1517
DB	575 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly	594
QY	1518 CGTGGGAATCGACATGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1577
DB	595 ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp	614
QY	1578 ATCACTTACATGTAAGACACATTCACAAAGCGTCAATGAAGGGGACACATGCTGAAC	1637
DB	615 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn	634
QY	1638 TGTACATGCTTCGCTCAGGGTGGGAGTGTGATCCGTCGACCAATGCCAG	1697
DB	635 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln	654
QY	1698 GATTACAGACCTGGGACGCTTTTATCAAAATTCGAGATTCATGGGAGAGTATGTGATGCT	1757
DB	655 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrValHisGly	674
QY	1758 GTCAGATACAGTGTCTGCTATGCGCGTGGCATTTGGGAGTGGCATTCGCAACTTCA	1817
DB	675 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu	694
QY	1818 CAGACCTATCAAGCTCAAGTGTCTGTCGCAAGTATTTATCAGTACGACTCCGAGTCCAG	1877
DB	695 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln	714
QY	1878 CCCAACTCCCAACCCCATCCAGTGAATGCACACAGCCATCTCACAATTTCCAAAGTACATT	1937
DB	715 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle	734
QY	1938 CTCAGTGGAGACTGTGAGTATCCACCCAGAAACCTTGGATAC	1982
DB	735 LeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr	749

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 11, 2005, 06:14:58 ; Search time 42 Seconds
(without alignments)
9837.019 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 4096
Sequence: 1 caaactgtggaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2 1/USPTO spool h/US09581651/runat 07112005 092114 27555/app query.fasta_1.2311
-DB=PIR -QFWT=fastan -SUPFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09581651 @CGN 1 1 55 @runat 07112005 092114 27555 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_79.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Length	ID	Description	
1	3579.5	87.4	2386	1 FNHU	fibronectin precu
2	3354.5	81.9	2477	2 S14428	fibronectin precu
3	3328.5	81.3	2265	1 FNBO	fibronectin - bovi
4	2888	70.5	2481	2 A43908	fibronectin - Afri
5	961	23.5	190	2 I51279	fibronectin - east
6	306	7.5	708	2 JC4364	gelatinase B (EC 3
7	301	7.3	662	2 S70355	gelatinase A (EC 3
8	300	7.3	662	2 A42496	gelatinase A (EC 3
9	300	7.3	662	2 S34780	gelatinase A (EC 3
10	300	7.3	707	1 A53796	gelatinase B (EC 3
11	300	7.3	708	2 S62907	gelatinase B (EC 3
12	298	7.3	663	1 S46492	gelatinase A (EC 3
13	297	7.3	660	1 A28153	gelatinase A (EC 3
14	294.5	7.2	712	1 I46031	gelatinase B (EC 3

15	294.5	7.2	730	1	152580	gelatinase B (EC 3
16	294.5	7.2	730	1	JC1456	gelatinase B (EC 3
17	285.5	7.0	707	1	A34458	gelatinase B (EC 3
18	228.5	5.6	1020	2	A29355	fibronectin - chic
19	211	5.2	1455	1	A48925	mannose receptor p
20	207	5.1	1456	1	A36563	mannose receptor p
C 21	206.5	5.2	1680	2	A43434	furin (EC 3.4.21.7
22	204	5.0	5376	2	T42215	zonadhesin - mouse
23	203.5	5.0	722	2	I48324	DELTA-like 1 - mou
24	200	4.9	1458	1	A49707	phospholipase A2 r
25	195.5	4.8	1479	2	T42710	mannose receptor,
26	188.5	4.6	1584	2	T22674	hypothetical prote
27	185.5	4.5	1326	2	B56395	secretory phosphol
28	185.5	4.5	1465	2	A56395	secretory phosphol
29	185.5	4.5	2531	2	S18188	notch protein homo
C 30	184	4.6	1299	2	T43251	furin (EC 3.4.21.7
31	183	4.5	473	2	A56175	adhesive plaque pr
32	182	4.4	2555	2	A40043	notch protein homo
33	180.5	4.4	2139	2	A35672	crumbs protein - f
34	179.5	4.4	728	2	I50719	C-Delta-1 - chicke
35	179.5	4.3	2531	2	A46019	phospholipase A2 r
36	177.5	4.3	3672	2	T23433	hypothetical prote
37	176.5	4.3	3704	2	T37316	probable laminin a
38	176.5	4.3	2437	2	S42612	transmembrane prot
C 39	176	4.3	2437	2	S42612	transmembrane prot
40	175.5	4.4	1737	2	T00209	MEGF8 protein - hu
41	174.5	4.3	1487	2	S48719	phospholipase-A(2)
42	174.5	4.3	2703	1	A24420	notch protein - fr
43	174	4.2	2543	2	T31687	surface antigen - p
44	173	4.2	884	2	T18649	hypothetical prote
45	171.5	4.2	3707	2	S18252	heparan sulfate pr

ALIGNMENTS

RESULT 1

FNHU

fibronectin precursor [validated] - human
N;Alternate names: fibronectin splice form ED-A
C;Species: Homo sapiens (man)
C;Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004
C;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22
R;Dean, D.C.; Bowie, C.L.; Bourgeois, S.
Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.
A;Reference number: A26460; MUID:87175578; PMID:3031656
A;Accession: A26460
A;Molecule type: DNA
A;Residues: 1-49 <DEA>
A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAAS337

R;Oldberg, A.; Ruoslahti, E.
J. Biol. Chem. 261, 2113-2116, 1986
A;Title: Evolution of the fibronectin gene.
A;Reference number: A26284; MUID:86111901; PMID:3003095
A;Accession: A26284
A;Molecule type: DNA
A;Residues: 1447-1540 <OLD>
A;Cross-references: GB:M12549; NID:g182688
A;Note: the authors translated the codon TTC for residue 1494 as Glu
R;Paolella, G.; Henschliffe, C.; Sebastio, G.; Baralle, F.E.
Nucleic Acids Res. 16, 3545-3557, 1988
A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B
A;Reference number: S00848; MUID:88233940; PMID:3375063
A;Accession: S03917
A;Molecule type: DNA
A;Residues: 1594-1767, 'V', 1769-1783 <PAO>
A;Cross-references: EMBL:X07718; NID:g31402
A;Note: the authors translated the codon AAC for residue 1631 as Asp
R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
FEBS Lett. 207, 287-291, 1986
A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene.
A;Reference number: A24854; MUID:87030925; PMID:3770201
A;Accession: A24854

A;Molecule type: DNA
A;Residues: 1992-2147 <VTB>
A;Cross-references: GB:X04530; NID:g31436
R;Gutman, A.; Yamada, K.M.; Kornblitt, A.
FEBS Lett. 207, 145-148, 1986
A;Title: Human fibronectin is synthesized as a pre-propolypeptide.
A;Reference number: A24476; MUID:87030890; PMID:3770189
A;Accession: A24476
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-14,'Q',16-38 <GUT>
R;Kornblitt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
EMBO J. 4, 1755-1759, 1985
A;Title: Primary structure of human fibronectin: differential splicing may generate at 1
A;Reference number: A91008; MUID:85284965; PMID:2992939
A;Accession: A91008
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 32-1344,1346-2080;2112-2386 <KOR>
A;Cross-references: GB:X02761
R;Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptide
A;Reference number: A93529; MUID:84272258; PMID:6462919
A;Accession: A93529
A;Molecule type: mRNA
A;Residues: 973-2080;2112-2386 <KO2>
A;Cross-references: GB:X00739
R;Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
A;Reference number: A21011; MUID:83290929; PMID:6688418
A;Accession: A21011
A;Molecule type: mRNA
A;Residues: 1434-1537 <OL2>
A;Cross-references: GB:K00055; NID:g182680; PIDN:AAA52459.1; PID:g182683
R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
A;Reference number: A90495; MUID:85280409; PMID:2992573
A;Accession: A90495
A;Molecule type: mRNA
A;Residues: 1594-2386 <BER>
A;Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697
R;Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A;Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A;Reference number: A22245; MUID:85231203; PMID:2989004
A;Accession: A22245
A;Molecule type: mRNA
A;Residues: 1948-2067 <UMB>
A;Cross-references: GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:g182706
A;Accession: B22245
A;Molecule type: mRNA
A;Residues: 1975-1991;2017-2039 <UM2>
A;Cross-references: GB:M27590
R;Sekiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A;Title: Human liver fibronectin complementary DNAs: identification of two different mes
A;Reference number: I52394; MUID:87026578; PMID:3021206
A;Accession: I65273
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1978-1990,2016-2018,'N',2020-2081,2113-2127 <SBK>
A;Cross-references: GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:g182704
R;Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A;Reference number: A21165; MUID:83221567; PMID:6304699
A;Accession: A21165
A;Molecule type: mRNA
A;Residues: 2291-2386 <KO3>
A;Cross-references: GB:K00799; NID:g182681; PIDN:AAA52460.1; PID:g182684

R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A;Title: Primary structure of human plasma fibronectin.
A;Reference number: A92398; MUID:84032463; PMID:6630202
A;Accession: A92398
A;Molecule type: protein
A;Residues: 32-47,'C',49-51,'S',53-72,'A',74-290 <GARI>
R;Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A;Title: Further characterization of the binding of fibronectin to gelatin reveals the p
A;Reference number: S34791; MUID:93312001; PMID:8323285
A;Accession: S34791
A;Molecule type: protein
A;Residues: 291-300;551-560 <GAR2>
R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A;Reference number: A60904; MUID:87019725; PMID:3532418
A;Accession: A60904
A;Molecule type: protein
A;Residues: 293-301 <GRI>
R;Calaycay, J.; Fand, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl
A;Reference number: A23901; MUID:86008277; PMID:3900070
A;Accession: A23901
A;Molecule type: protein
A;Residues: 616-677,'Q',679-703,'PT' <CAL>
R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
A;Reference number: A92386; MUID:82265604; PMID:7050098
A;Accession: A92386
A;Molecule type: protein
A;Residues: 1441-1548 <PIE>
A;Note: residues 1524-1527 are responsible for the cell-binding activity
R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A;Reference number: A32517; MUID:87241275; PMID:3593230
A;Accession: A32517
A;Molecule type: protein
A;Residues: 1589-1630,'T',1722-2058 <GAR3>
R;Tressel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A;Title: Human plasma fibronectin. Demonstration of structural differences between the A
A;Reference number: S14357; MUID:91190085; PMID:2012601
A;Accession: S14357
A;Molecule type: protein
A;Residues: 1614-1630,'T',1722-2081,2113-2244 <TRE>
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A;Reference number: A23891; MUID:85261459; PMID:4019516
A;Accession: A23891
A;Molecule type: protein
A;Residues: 2071-2080;2112-2356 <GAR4>
C;Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
ation, and transformation.
C;Genetics:
A;Gene: GDB:FNI
A;Cross-references: GDB:I19135; OMIM:135600
A;Map position: 2q34-2q34
A;Intons: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-31/Domain: propeptide #status predicted <PRO>
F;32-2386/Product: fibronectin #status experimental <MAT>
F;52-272/Domain: fibrin and heparin binding <FHB>
F;52-87/Domain: fibronectin type I repeat homology <1F1>
F;97-135/Domain: fibronectin type I repeat homology <1F2>

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 11, 2005, 07:54:49 ; Search time 171.5 Seconds
(without alignments)
10476.095 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 4096

Sequence: 1 caaacttggtggcaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 3735758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPRO_spool_h/US09581651/runat_07112005_092117_27656/app_query.fasta_1.2311
-DB=Published Applications_AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pet -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -WAP=US09581651@cgn1.1.161 @runat_07112005_092117_27656
-NCPU=6 -ICPU=3 -NO_MMAP -LARGESQ=0 -NEG SCORES=0 -WAIT -DSBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAD=10 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US16_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3664	89.5	642	16	US-10-741-601-354
2	3664	89.5	642	17	US-10-741-600-1066
3	3646.5	89.0	657	16	US-10-741-601-359
4	3646.5	89.0	657	17	US-10-741-600-1072
5	3599.5	87.9	2220	15	US-10-236-392-4
6	3599.5	87.9	2355	15	US-10-144-194A-104
7	3599.5	87.9	2355	15	US-10-447-161-3
8	3599.5	87.9	2355	16	US-10-734-564-94
9	3599.5	87.9	2355	16	US-10-491-566-104
10	3599.5	87.9	2355	17	US-10-852-335A-147
11	3599.5	87.9	2355	18	US-10-287-436A-436
12	3599.5	87.9	2355	18	US-10-287-436A-1137
13	3599.5	87.9	2355	20	US-11-040-130-28
14	3599.5	87.9	2386	16	US-10-618-281-32
15	3599.5	87.9	2474	18	US-10-450-763-52638
16	3595.5	87.8	2386	10	US-09-961-403-1
17	3595.5	87.8	2386	16	US-10-788-792-206
18	3595.5	87.8	2386	17	US-10-868-577A-59
19	3595.5	87.8	2386	18	US-10-485-758-4
20	3595.5	87.8	2386	18	US-10-485-758-9
21	3592.5	87.7	984	16	US-10-741-601-356
22	3592.5	87.7	984	17	US-10-741-600-1069
23	3592.5	87.7	2296	16	US-10-741-601-363
24	3592.5	87.7	2296	17	US-10-741-600-1075
25	3592.5	87.7	2355	16	US-10-741-601-357
26	3592.5	87.7	2355	16	US-10-741-601-366
27	3592.5	87.7	2355	17	US-10-741-600-1087
28	3592.5	87.7	2355	17	US-10-741-600-1078
29	3592.5	87.7	2386	16	US-10-741-601-360
30	3592.5	87.7	2386	17	US-10-741-600-1071
31	3448.5	84.2	2328	14	US-10-171-311-64
32	3448.5	84.2	2328	15	US-10-236-031B-70
33	3448.5	84.2	2328	15	US-10-374-979-98
34	3448.5	84.2	2328	15	US-10-182-936A-98
35	3448.5	84.2	2328	16	US-10-477-238A-677
36	3448.5	84.2	2328	16	US-10-680-287A-677
37	3448.5	84.2	2328	17	US-10-477-173-677
38	3439.5	84.0	2355	15	US-10-360-101-235
39	3424.5	83.6	2320	14	US-10-279-733-8
40	3410.5	83.3	2320	15	US-10-236-392-2
41	3296.5	80.5	1173	18	US-10-450-763-52634
42	2235	54.6	463	15	US-10-144-194A-52
43	2235	54.6	463	16	US-10-491-566-52
44	1859.5	45.4	343	9	US-09-934-706-1
45	1859.5	45.4	400	9	US-09-934-706-5

ALIGNMENTS

RESULT 1
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Alignment Scores:

Pred. No.:	2, 04e-280	Length:	642
Score:	3664.00	Matches:	639
Percent Similarity:	99.53%	Conservative:	0
Best Local Similarity:	99.53%	Mismatches:	3
Query Match:	89.45%	Indels:	0
DB:	16	Gaps:	0

US-09-581-651D-3 (1-2147) x US-10-741-601-354 (1-642)

Qy	57	ATGCTTAGGGTCCGGGCGCCGGCTGCTGTGCTGGCCGCTCCAGTGCCTGGGACACG	116
Db	1	MetLeuArgGlyProGlyProGlyLeuLeuLeuLeuAlaValLeuCysLeuGlyThrAla	20
Qy	117	GTGCGCTCCAGGAGCCTCGAAGAGCAAGAGGAGGCTCAGCAATGGTTCACGCCCG	176
Db	21	ValProSerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGln	40
Qy	177	TCCCGGTGGCTGTCAGTCAAGCAAGCCGGTGTATGATCAATGGAAAAACACTATCAG	236
Db	41	SerProValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGln	60
Qy	237	ATAAATCAACAGTGGAGCGGACCTACCTAGGCAATGCGTGGTGTGTTGTTATGGA	296
Db	61	IleAsnGlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGly	80
Qy	297	GGAGCCGAGGTTTAACTCGGAGAGTAACCTGAGCTGAGAGACTTCTTTCACACG	356
Db	81	GlySerArgGlyPheAsnCysGluSerLysProGluAlaGluGlnThrCysPheAspLys	100
Qy	357	TACACTGGGAACACTTACCGAGTGGTGACACTATGAGCGCTCAAGAGACTTCCATGATC	416
Db	101	TyrThrGlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIle	120
Qy	417	TGGGACTGTACTGTGATCGGGCTGGCGAGGAGAGTAAGCTGTACCATCGCAACCGC	476
Db	121	TrpAspCysThrCysIleGlyAlaGlyArgGlyArgGlySerCysThrIleAlaAsnArg	140
Qy	477	TGCCATGAAGGGGTGAGTCTCAAGATTGGTGCACCTGGAGAGACCATGAGACT	536
Db	141	CysHisGluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgProHisGluThr	160
Qy	537	GGTGGTGTACATGTAGAGTGTGTGCTTGGTAATGGAAGAGAGATGGACCTGCAAG	596
Db	161	GlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLys	180
Qy	597	CCATAGTGTAGAAAGTGTGATCATGCTGTGGGACTTCTATGTGTGGAGAAACG	656
Db	181	ProIleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr	200
Qy	657	TGGGAGAACCTTACCAAGGCTGGATGATGTAGATTGTACTTGTGGGAGGAGGCAGC	716
Db	201	TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer	220
Qy	717	GGAGCGCATCAGTGTGACCTTCTAGAAATAGATCAACAGATCAGGACACAGGACATCCTAT	776
Db	221	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	240
Qy	777	AGAATTGGAGACACTGGAGCAGGAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	836
Db	241	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys	260
Qy	837	ACAGCGAACGGCGAGAGAGTGGAGTGTGAGAGCACACCTCTGTGAGACCATCG	896
Db	261	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	280
Qy	897	AGCGATCTGGCCCTTACCGATGTTGCTGCGAGCTGTTTACCAACCGCAGCCTCACCCC	956
Db	281	SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro	300
Qy	957	GAGCTCTCTCCATGTGGCCACTGTGTACAGACAGTGGTGGTGTCTCTGTGGGATG	1016
Db	301	GlnProProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet	320

Qy	1017	CAGTGGCTGAAGACACACAAAGAAATAGCAATGCTTTGACGCTGCTGGGCAACGAGTC	1076
Db	321	GlnTrpLeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyVal	340
Qy	1077	AGTGGCCAAAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGAGAGCCATGT	1136
Db	341	SerCysGlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCys	360
Qy	1137	GTCTTACCATTCACCTACACAGCAGCAGCAGCACTTCCGAATATTGACGAGAC	1196
Db	361	ValLeuProPheThrTyrAsnGlyArgThrCysSerThrThrSerAsnTyrGluGlnAsp	380
Qy	1197	CAGAAATACCTCTTCTGCACAGACACACTGTTTGGTTCAGACTCCAGAGAGGAAATCC	1256
Db	381	GlnLysTyrSerPheCysThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSer	400
Qy	1257	AATGGTCTGTGTCCTTCCCTTCTTATACAAACCAACCAATTAACATGATGACT	1316
Db	401	AsnGlyAlaLeuCysHisPheProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThr	420
Qy	1317	TCTGAGGCGCAGAGAGACACATGAGTGGTGGGACCAACACAGAACTATGATGCCGAC	1376
Db	421	SerGluGlyArgArgAspAsnMetLysTyrCysGlyThrGlnAsnTyrAspAlaAsp	440
Qy	1377	CAGAAATTTGGTTCCTCCCATGCTGGTGGGAAATCTGCACAAACCAATGAAGGG	1436
Db	441	GlnLysPheGlyPheCysProMetAlaAlaHisGluGluIleCysThrThrAsnGluGly	460
Qy	1437	GTCAATGATCCGCAATGGAGATCAGTGGGATAAGCAGCATGACATGGGTGCATGATGAG	1496
Db	461	ValMetTyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArg	480
Qy	1497	TGCACGTGTGTGGGAATGCTGGGGAATGGATGATGCTGCTACTGTCACCTCGA	1556
Db	481	CysThrCysValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArg	500
Qy	1557	GATCAGTGCATGTTGTGATGACATCATTACATGTGAACGACACATTCACCAACGCTCAT	1616
Db	501	AspGlnCysIleValAspAspIleThrTyrAsnValAsnAspThrPheHisLysArgHis	520
Qy	1617	GAACAGGGGCACATGCTGAACATGTATCATGCTTCGGTCAGGGTCGGGCGAGTGAAGTGT	1676
Db	521	GluGluGlyHisMetLeuAsnCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCys	540
Qy	1677	GATCCCTCGACCAATGCCAGATTTCAGAGACTGGGACGTTTATCAAAATGGAGATTCA	1736
Db	541	AspProValAspGlnCysGlnAspSerGluThrGlyThrPheTyrGlnIleGlyAspSer	560
Qy	1737	TGGGAGAAGTATGTGTCATGCTGATACAGATACCAAGTGTCTGCTATGGCCGTGGCATTTGG	1796
Db	561	TrpGluLysTyrValHisGlyValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGly	580
Qy	1797	GAGTGGCATTTGCCAACCTTTACAGACCTTATCAAGCTTCAAGTGTCTGCTGTCGAGTATTT	1856
Db	581	GluTrpHisCysGlnProLeuGlnThrTyrProSerSerSerGlyProValGluValPhe	600
Qy	1857	ATCACTGAGACTCCGAGTCAGCCCACTCCACCCCATCCAGTGGAGTCCACCCAGGCA	1916
Db	601	IleThrGluThrProSerGlnProAsnSerHisProIleGlnThrAsnAlaProGlnPro	620
Qy	1917	TCTCACATTTCCAAAGTACATTTCTCAGGTGGAGACTGTGAGTATCCCAACCCAGAACCTT	1976
Db	621	SerHisIleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeu	640
Qy	1977	GGATAC 1982	
Db	641	GlyTyr 642	

RESULT 2
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:

12 298 7.3 663 4 US-09-194-468A-30 Sequence 30, 1

Result No.	Score	Query Match	Length	DB	ID	Description
------------	-------	-------------	--------	----	----	-------------

Db 1 MetLeuArgGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAla 20
Qy 117 GTGCTCCAGGAGGCTCGAAGAGAGAGGAGGCTGAGAAATGGTTTCAGCCCGAG 176
Db 21 ValProSerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGln 40
Qy 177 TCCCGGTGGGTGTCAGTCAAGCAAGCCCGTTGTTATGACAAATGGAACACATATCAG 236
Db 41 SerProValAlaValSerGlnSerLysProGlyCysThrAspAenGlyLysHisThrGln 60
Qy 237 ATAAATCAACAGTGGGAGCGACCTTACCTAGGCAATGCTGGTTGTTGTTATGGA 296
Db 61 IleAsnGlnInTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysThrGly 80
Qy 297 GGAAGCGAGGTTTAACTCGAGAGTAAACCTGAAGCTGAAGAGACTGCTTTGACAAAG 356
Db 81 GlySerArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLys 100
Qy 357 TACACTGGACACTTACCGAGTGGTGACACTTATGAGCGTCTAAAGACTCCATGATC 416
Db 101 TyrThrGlyAenThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIle 120
Qy 417 TGGACTGTACTGCTCGGCGTGGGAGGAGATAAGCTGTACCATCGCAACCGC 476
Db 121 TrpAspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArg 140
Qy 477 TGCCATGAAGGGGTCACTCTACAGATTGGTGACACCTCGAGAGACCATGAGACT 536
Db 141 CysHisGluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThr 160
Qy 537 GGTGGTACATGTTAGTGTGTCTGCTGTAATGGAAGGAGATGACCTGCAAG 596
Db 161 GlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLys 180
Qy 597 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCTATGTGGTGGGAGAACG 656
Db 181 ProIleAlaGlyLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 200
Qy 657 TGGAGAAAGCCCTACCAAGCTGGATGATGTAGATTGTCTGCTGGGAGAGGAGC 716
Db 201 TrpGlyLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 220
Qy 717 GGAGCATCACTGTCACCTTAGAATAAGTACACGATCAGGACACAAAGGACATCCTAT 776
Db 221 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 240
Qy 777 AGAATTGGAGACACCTGGAGCAAGGATATCGAGGAAACCTGCTCCAGTGCATCTGC 836
Db 241 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 260
Qy 837 ACAGGCAACGGCCGAGGAGTGGAAAGTGTGAGAGGCAACCTCTGTGCAGACCATCG 896
Db 261 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 280
Qy 897 AGCGCATCTGSCCCTTACCGCATCTCGTGAGCTGTTTACCAACCCAGCCTCACCCC 956
Db 281 SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro 300
Qy 957 CAGCCTCTCTCTATGGCCACTGTGTCACAGACAGTGGTGGTCTACTCTGTGGGATG 1016
Db 301 GlnProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet 320
Qy 1017 CAGTGGCTGAGACACAGAAATAGCAATGCTTTGACGTGCTGGGCAAGGAGTC 1076
Db 321 GlnTrpLeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyVal 340
Qy 1077 AGCTCCCAAGACACAGCTGTACCCAGACTTACGTTGGCAACTCAATGGAGAGCCATGT 1136
Db 341 SerCysGlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCys 360
Qy 1137 GTCTTACCATTACCTACCAACGACAGGAGC----- 1166
Db 361 ValLeuProPheThrTyrAsnGlyArgThrPheTyrSerCysThrThrGluGlyArgGln 380

Qy 1167 -----GACAGCACAACTTCGAATTATGAGCAGGACGACGAAATACTCTTTT 1211
Db 381 AspGlyHisLeuTrpCysSerThrThrSerAsnTyrGlnGlnAspGlnLysTyrSerPhe 400
Qy 1212 TGCAAGACACACACTGTTTTGGTTTCAGACTCGAGGAGAAATTCGAATGTGCTTGTGC 1271
Db 401 CysThrAspHisThrValLeuValGlnThrGlnGlyGlyAsnSerAsnGlyAlaLeuCys 420
Qy 1272 CACTTCCCTTCTATACAAACCAACCAATTACACTGATTGCACCTTCTGAGGCGCAGAGA 1331
Db 421 HisPheProPheLeuTyrAsnAenHisAsnTyrThrAspCysThrSerGluGlyArgArg 440
Qy 1332 GACAACTGAAGTGGTGGGACACACAGAACTATGATGCCACCAAGAGTTTGGGTTTC 1391
Db 441 AspAenMetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPhe 460
Qy 1392 TGCCCCATGGCTGCCACGAGGAAATCTGCACAACTCAATCAAGGGGTTCATGTACCGCAT 1451
Db 461 CysProMetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIle 480
Qy 1452 GGAGATCAGTGGGATTAAGCAGCATGACATGGGTACATGATGAGGTGCAGCTGTGTTGGG 1511
Db 481 GlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGly 500
Qy 1512 AATGGTCTGGGAAATGGACATGCATTGCTCTACTCGCAGCTTCGAGATCAGTGCATTGTT 1571
Db 501 AsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleVal 520
Qy 1572 GATGACATCACTTACAAATGTGAACGACACATTCACCAAGCGTCATGAAGGGGCGACATG 1631
Db 521 AspAspIleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGlyHisMet 540
Qy 1632 CTGAACCTGTACATCTTCGGTTCAGGCTCGGCGCAGGTGGAAAGTGTGATCCGTCGACCA 1691
Db 541 LeuAsnCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGln 560
Qy 1692 TGCCAGGATTCAGAGACTGGGAGCTTTTATCAATTCGAGATTCATGGGAGAGTATGTG 1751
Db 561 CysGlnAspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrVal 580
Qy 1752 CATGGTTCAGATACCACTGCTACTGCTATGGCGCTGGCATTTGGGAGTGGCATGCCAA 1811
Db 581 HisGlyValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGln 600
Qy 1812 CTTTACAGACCTATCCAAGCTCAAGTGCCTGCTCGAAGTATTTATCCTGAGACTCCG 1871
Db 601 ProLeuGlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrPro 620
Qy 1872 AGTCAGCCCACTCCCAACCCCATCCAGTGAATGCACACAGCCATCTCACATTTTCCAAG 1931
Db 621 SerGlnProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLys 640
Qy 1932 TACATTTCTCAGGTGGAGACCT 1952
Db 641 TyrIleLeuArgTrpArgPro 647
RESULT 2
US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 11, 2005, 03:51:06 ; Search time 163.5 seconds
(without alignments)
10157.487 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 4096

Sequence: 1 caacttggtggcaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgm2_1/USPTO.spool_b/US09581651/runat_07112005_092112_27535/app_query.fasta_1.2311
-DB=A_Geneseq -QFWT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CGN 1.1 163 @runat_07112005_092112_27535 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3781	92.3	660	2 AAY28901	Aay28901 Human mig
2	3693.5	90.2	2182	8 ADR66462	Adr66462 Human pro
3	3693.5	90.2	2182	8 ADR66120	Adr66120 Human pro
4	3680.5	89.9	720	2 AAY28914	Aay28914 Fibronect
5	3680	89.8	642	8 ADR67316	Adr67316 Human bla
6	3680	89.8	642	8 ADS17489	Adsl7489 Amino aci
7	3680	89.8	642	8 ADR97658	Adr97658 Human fib
8	3664	89.5	642	8 ADQ39403	Adq39403 Human myo
9	3658	89.3	2447	4 AAM40434	Aam40434 Human pol
10	3658	89.3	2447	4 AAM40435	Aam40435 Human pol

11	3658	89.3	2447	4 AAM40433	Aam40433 Human pol
12	3658	89.3	2447	4 AAM40432	Aam40432 Human pol
13	3646.5	89.0	657	8 ADQ39409	Adq39409 Human myo
14	3599.5	87.9	2220	6 ABO01289	Abo01289 Human pro
15	3599.5	87.9	2266	6 ABR40124	Abr40124 Human cel
16	3599.5	87.9	2355	6 ABR58335	Abr58335 NM_00202
17	3599.5	87.9	2355	7 ADP65196	Adp65196 Human fib
18	3599.5	87.9	2355	8 ADG89560	Adg89560 Human fib
19	3599.5	87.9	2355	8 ADO55175	Ado55175 Protein #
20	3599.5	87.9	2355	8 ADQ36085	Adq36085 Fibronect
21	3599.5	87.9	2355	8 ADQ29668	Adq29668 Human col
22	3599.5	87.9	2355	8 ADR67315	Adr67315 Human bla
23	3599.5	87.9	2386	8 ADO55174	Ado55174 Protein #
24	3599.5	87.9	2474	4 ABG22279	Abg22279 Novel hum
25	3595.5	87.8	2386	5 AAO17353	Aao17353 Human fib
26	3595.5	87.8	2386	6 ABR81866	Abr81866 Human fib
27	3595.5	87.8	2386	7 ADD18770	Add18770 Human dia
28	3595.5	87.8	2386	7 ADE63324	Ade63324 Human pro
29	3595.5	87.8	2386	8 ADR99200	Adr99200 Fibronect
30	3593.5	87.7	2446	3 AAB50377	Aab50377 Human fib
31	3592.5	87.7	984	8 ADQ39406	Adq39406 Human myo
32	3592.5	87.7	2265	4 AAM38647	Aam38647 Human pol
33	3592.5	87.7	2296	8 ADQ39412	Adq39412 Human myo
34	3592.5	87.7	2330	4 AAM38646	Aam38646 Human pol
35	3592.5	87.7	2355	4 AAM38649	Aam38649 Human pol
36	3592.5	87.7	2355	8 ADQ39415	Adq39415 Human myo
37	3592.5	87.7	2355	8 ADQ39404	Adq39404 Human myo
38	3592.5	87.7	2386	4 AAM38648	Aam38648 Human pol
39	3592.5	87.7	2386	8 ADQ39408	Adq39408 Human myo
40	3586.5	87.6	2477	2 AAM99595	Aam99595 Human fib
41	3584.5	87.5	2220	8 ADN95950	Adn95950 Human NOV
42	3579.5	87.4	2386	2 AAW63171	Aaw63171 Amino aci
43	3578.5	87.4	2446	2 AAR60021	Aar60021 Fibronoge
44	3546	86.6	2352	8 ADK00410	Adk00410 Fibronect
45	3448.5	84.2	2328	4 AAG68182	Aag68182 Fibronect

ALIGNMENTS

RESULT 1

AAY28901
ID AAY28901 standard; protein; 660 AA.

XX AAY28901;

XX 21-SEP-1999 (first entry)

XX Human migration stimulating factor (MSF) 1-alpha protein.

XX Migration stimulatory factor; MSF; cell migration; modulation; human; wound healing; scarring; MSF1-alpha.

XX Homo sapiens.

XX WO9931233-A1.

XX 24-JUN-1999.

XX 15-DEC-1998; 98WO-GB003766.

XX 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI; 1999-430039/36.

XX N-PSDE; AAX81299.

XX Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX Claim 1; Page 53; 86pp; English.

PS

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 202.745 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-1
Perfect score: 3781
Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILRWRPVPSPRNLGY 660

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3732.5	98.7	749	2 Q7Z391	Q7Z391 homo sapien
2	3686.5	97.5	1103	2 Q6M2F4	Q6M2F4 homo sapien
3	3685.5	97.5	2357	2 Q68D74	Q68D74 homo sapien
4	3682.5	97.4	2444	2 Q6N025	Q6N025 homo sapien
5	3595.5	95.1	2386	1 F1NC_HUMAN	P02751 homo sapien
6	3592.5	95.0	2296	2 Q6N0A6	Q6N0A6 homo sapien
7	3592.5	95.0	2477	2 Q6MZU5	Q6MZU5 homo sapien
8	3588.5	94.9	2267	2 Q68DP8	Q68DP8 homo sapien
9	3567.5	94.4	2240	2 Q68DP9	Q68DP9 homo sapien
10	3393.5	89.8	2477	1 F1NC_MOUSE	P11276 mus musculus
11	3354.5	88.7	2477	1 F1NC_RAT	P04937 rattus norv
12	3328.5	88.0	2265	1 F1NC_BOVIN	P07589 bos taurus
13	2888	76.4	2481	2 Q6GQAS	Q6GQAS xenopus lae
14	2878	76.1	2481	1 F1NC_XENLA	Q91740 xenopus lae
15	2513.5	66.5	922	2 Q93405	Q93405 brachydanio
16	2513.5	66.5	2478	2 Q93406	Q93406 brachydanio
17	2328	61.6	2408	2 Q6JAN2	Q6JAN2 brachydanio
18	1909.5	50.5	2193	2 Q6CZM7	Q6CZM7 homo sapien
19	1496	39.6	296	2 Q8C6J7	Q8C6J7 mus musculus
20	961	25.4	190	1 F1NC_NOTVI	Q91400 notophthalm
21	750	19.8	141	2 Q90XQ2	Q90XQ2 ambystoma m
22	541	14.3	215	2 Q6DD34	Q6DD34 xenopus lae
23	317.5	8.4	679	2 Q98856	Q98856 cynops pyrr
24	313.5	8.3	673	2 Q90YB3	Q90YB3 paralicthy
25	310.5	8.2	690	2 Q9PVM5	Q9PVM5 oryzias lat
26	307	8.1	670	2 Q6DF16	Q6DF16 xenopus tro
27	307	8.1	680	2 Q7T317	Q7T317 brachydanio
28	306	8.1	675	2 Q8QF06	Q8QF06 oncorhynch
29	302.5	8.0	810	2 Q8R3F3	Q8R3F3 mus musculus
30	302	8.0	671	2 Q6PF33	Q6PF33 xenopus lae
31	302	8.0	671	2 Q6W716	Q6W716 xenopus lae

32	301	8.0	655	2 Q9W635	Q9W635 oncorhynch
33	301	8.0	662	1 MM02_RABIT	P50757 oryctolagus
34	300.5	7.9	216	2 Q99KD0	Q99KD0 mus musculus
35	300.5	7.9	674	2 Q98TC6	Q98TC6 cyprinus ca
36	300	7.9	662	1 MM02_MOUSE	P33434 mus musculus
37	300	7.9	662	1 MM02_RAT	P33436 rattus norv
38	300	7.9	707	1 MM09_RABIT	P41246 oryctolagus
39	299.5	7.9	2146	2 Q68CX6	Q68CX6 homo sapien
40	298	7.9	412	2 Q71U44	Q71U44 bos taurus
41	298	7.9	663	1 MM02_CHICK	Q90611 gallus gall
42	297	7.9	660	1 MM02_HUMAN	P08253 homo sapien
43	296.5	7.8	1256	1 F1NC_CHICK	P11722 gallus gall
44	296	7.8	632	2 Q9N1P6	Q9N1P6 canis faml
45	295	7.8	211	2 Q7L553	Q7L553 homo sapien

ALIGNMENTS

RESULT 1
Q7Z391 Q7Z391 PRELIMINARY; PRT; 749 AA.
AC Q7Z391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloecher H., Bocher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -
DR HSP; Q96KP7; 1PBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006205; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 98.7%; Score 3732.5; DB 2; Length 749;
Best Local Similarity 97.2%; Pred. No. 2.6e-261;
Matches 656; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

Qy	1	NLVATCLPVRASLPHRLNMLRGPGGLLLAVQLGTAVPSTGASKSKRQAOQWVQOSP	60
Db	75	NLVATCLPVRASLPHRLNMLRGPGGLLLAVLCLGTAVPSTGASKSKRQAOQWVQOSP	134
Qy	61	VAVSQSGPGCVDNKGKHYQINQWERTYLGVALVCTCYGGSGRFGNCSKPEAEETCFDKYT	120
Db	135	VAVSQSGPGCVDNKGKHYQINQWERTYLGVALVCTCYGGSGRFGNCSKPEAEETCFDKYT	194
Qy	121	GNTYRVGDTYERPKDSMIWDTCTCIGAGRGISCTIANRCHGGQSYKIGDTRWRPHETGG	180
Db	195	GNTYRVGDTYERPKDSMIWDTCTCIGAGRGISCTIANRCHGGQSYKIGDTRWRPHETGG	254
Qy	181	YMLECVCLNGKGBWTCKPTAEKCFDHAAGTSYVVGITWETKPYQGMWVDCTCLGEGSGR	240

Db 255 YMLSCVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWKEPYQGMWMDCTCLGEGSGR 314
Qy 241 ITCTSRNRCDQDTRTSYRIGDTSKDNKGNLLQICITGNGRGEWKCERHTSVQTTSSG 300
Db 315 ITCTSRNRCDQDTRTSYRIGDTSKDNKGNLLQICITGNGRGEWKCERHTSVQTTSSG 374
Qy 301 SGPTFDVRAAYVQPPHPPPPYGHCVTDGVTYVSGWMLKTQGNKQMLCTCLGNGVSC 360
Db 375 SGPTFDVRAAYVQPPHPPPPYGHCVTDGVTYVSGWMLKTQGNKQMLCTCLGNGVSC 434
Qy 361 QETAVTQTYGNSNGEPCVLPFTYNDRT-----DSTTSNYEQDQKYSFCT 405
Db 435 QETAVTQTYGNSNEEPCVLPFTYNGRTFYSCITTEGRDGLWCSTTSNYEQDQKYSFCT 494
Qy 406 DHTVLVQTRGNSNGALCHFFPFLYNNHNYTCTSEGRDNNKWCCTTQNYDADQKFGFCP 465
Db 495 DHTVLVQTRGNSNGALCHFFPFLYNNHNYTCTSEGRDNNKWCCTTQNYDADQKFGFCP 554
Qy 466 MAHEEICTTNEGVMYRIGDQWDKQHDGMHMRCTCVGNGRGWETCIAYSQLRDQCI VDD 525
Db 555 MAHEEICTTNEGVMYRIGDQWDKQHDGMHMRCTCVGNGRGWETCIAYSQLRDQCI VDD 614
Qy 526 ITYNVNDTFHKEEGHMLNCTCFQGRGRWKCDPVDQCDSETGTFTFYIGDSWEKYVHG 585
Db 615 ITYNVNDTFHKEEGHMLNCTCFQGRGRWKCDPVDQCDSETGTFTFYIGDSWEKYVHG 674
Qy 586 VRYQCYCVRGIGGEWHCOPLQTPYSSSGPVEVFTETPSQNSHPIONNAPQSHISKYI 645
Db 675 VRYQCYCVRGIGGEWHCOPLQTPYSSSGPVEVFTETPSQNSHPIONNAPQSHISKYI 734
Qy 646 LRWRPVSIPPRNLGY 660
Db 735 LRWRPVSIPPRNLGY 749

RESULT 2
Q6MZP4
ID Q6MZP4 PRELIMINARY; PRT; 1103 AA.
AC Q6MZP4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219)
DE (Fragment)
GN Name=DKFZp686K139; Synonyms=DKFZp686F219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG The German Human cDNA Consortium;
RA Ansgore W., Krieger S., Regier T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amlid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BX649182; CAB46200.1; -;
DR EMBL; BX640802; CAB45885.1; -;
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR00562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.

DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 4.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82PEC4CAF634AD56 CRC64;
Query Match 97.5%; Score 3686.5; DB 2; Length 1103;
Best Local Similarity 97.3%; Pred. No. 8.2e-258;
Matches 647; Conservative 0; Mismatches 3; Indels 15; Gaps 1;
Qy 1 NLVATCLPVRASLPRLNMLRGPGLLLAVQCLGTAVPSTGASKRQAAQVQPSQ 60
Db 72 NLVATCLPVRASLPRLNMLRGPGLLLAVLCLGTAVPSTGASKRQAAQVQPSQ 131
Qy 61 VAVSQSPGCVYDNGKHQYQINQWERTYLGNAVCTCYGSGRGFNCEKPEABETCFDKYT 120
Db 132 VAVSQSPGCVYDNGKHQYQINQWERTYLGNAVCTCYGSGRGFNCEKPEABETCFDKYT 191
Qy 121 GNTYRVGDTYERPKDSIMWDCTCIGAGRGISCTIANRCHGGQSYKIGDTRRPHETGG 180
Db 192 GNTYRVGDTYERPKDSIMWDCTCIGAGRGISCTIANRCHGGQSYKIGDTRRPHETGG 251
Qy 181 YMLSCVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWKEPYQGMWMDCTCLGEGSGR 240
Db 252 YMLSCVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGTEWKEPYQGMWMDCTCLGEGSGR 311
Qy 241 ITCTSRNRCDQDTRTSYRIGDTSKDNKGNLLQICITGNGRGEWKCERHTSVQTTSSG 300
Db 312 ITCTSRNRCDQDTRTSYRIGDTSKDNKGNLLQICITGNGRGEWKCERHTSVQTTSSG 371
Qy 301 SGPTFDVRAAYVQPPHPPPPYGHCVTDGVTYVSGWMLKTQGNKQMLCTCLGNGVSC 360
Db 372 SGPTFDVRAAYVQPPHPPPPYGHCVTDGVTYVSGWMLKTQGNKQMLCTCLGNGVSC 431
Qy 361 QETAVTQTYGNSNGEPCVLPFTYNDRT-----DSTTSNYEQDQKYSFCT 405
Db 432 QETAVTQTYGNSNGEPCVLPFTYNGRTFYSCITTEGRDGLWCSTTSNYEQDQKYSFCT 491
Qy 406 DHTVLVQTRGNSNGALCHFFPFLYNNHNYTCTSEGRDNNKWCCTTQNYDADQKFGFCP 465
Db 492 DHTVLVQTRGNSNGALCHFFPFLYNNHNYTCTSEGRDNNKWCCTTQNYDADQKFGFCP 551
Qy 466 MAHEEICTTNEGVMYRIGDQWDKQHDGMHMRCTCVGNGRGWETCIAYSQLRDQCI VDD 525
Db 552 MAHEEICTTNEGVMYRIGDQWDKQHDGMHMRCTCVGNGRGWETCIAYSQLRDQCI VDD 611
Qy 526 ITYNVNDTFHKEEGHMLNCTCFQGRGRWKCDPVDQCDSETGTFTFYIGDSWEKYVHG 585
Db 612 ITYNVNDTFHKEEGHMLNCTCFQGRGRWKCDPVDQCDSETGTFTFYIGDSWEKYVHG 671
Qy 586 VRYQCYCVRGIGGEWHCOPLQTPYSSSGPVEVFTETPSQNSHPIONNAPQSHISKYI 645
Db 672 VRYQCYCVRGIGGEWHCOPLQTPYSSSGPVEVFTETPSQNSHPIONNAPQSHISKYI 731
Qy 646 LRWRP 650
Db 732 LRWRP 736

RESULT 3
Q68DT4
ID Q68DT4 PRELIMINARY; PRT; 2357 AA.
AC Q68DT4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686F10164.
GN Name=DKFZp686F10164;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 50.0392 Seconds
(without alignments)
1269.066 Million cell updates/sec

Title: US-09-581-651D-1
Perfect score: 3781
Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILWRPVSIPPRNLGY 660

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3579.5	94.7	2386	1 FNHU	fibronectin precu
2	3354.5	88.7	2477	2 S14428	fibronectin precu
3	3328.5	88.0	2265	1 FNBO	fibronectin - bovi
4	2888	76.4	2481	2 A43908	fibronectin - Afri
5	961	25.4	190	2 I51279	fibronectin - east
6	302.5	8.0	708	2 JC4354	gelatinase B (EC 3
7	301	8.0	662	2 S70365	gelatinase A (EC 3
8	300	7.9	662	2 A42496	gelatinase A (EC 3
9	300	7.9	662	2 S34780	gelatinase A (EC 3
10	300	7.9	707	1 A53796	gelatinase B (EC 3
11	298	7.9	663	1 S46492	gelatinase A (EC 3
12	297	7.9	660	1 A28153	gelatinase A (EC 3
13	294.5	7.8	708	2 S62907	gelatinase B (EC 3
14	294.5	7.8	712	1 I46031	gelatinase B (EC 3
15	294.5	7.8	730	1 I52580	gelatinase B (EC 3
16	294.5	7.8	730	2 JC1456	gelatinase B (EC 3
17	285.5	7.6	707	1 A34458	gelatinase B (EC 3
18	228.5	6.0	1020	2 A29355	fibronectin - chic
19	208	5.5	1455	1 A48925	mannose receptor p
20	199.5	5.3	1456	1 A36563	mannose receptor p
21	199	5.3	5376	2 T42215	mannose receptor,
22	190.5	5.0	1479	2 T42210	mannose receptor,
23	188.5	5.0	1584	2 T22674	hypothetical prote
24	188	5.0	1458	1 A49707	phospholipase A2 r
25	185.5	4.9	1326	2 B56395	secretory phosphol
26	185.5	4.9	1465	2 A56395	secretory phosphol
27	182	4.8	722	2 I48324	DELTA-like 1 - mou
28	181.5	4.8	473	2 A56175	adhesive plaque pr
29	179.5	4.7	1463	2 A53210	phospholipase A2 r

RESULT 1

FNHU

fibronectin precursor [validated] - human
N;Alternate names: fibronectin splice form ED-A

C;Species: Homo sapiens (man)

C;Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004

C;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22

R;Dean, D.C.; Bowlus, C.L.; Bourgeois, S

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A;Reference number: A26460; MUID:87175578; PMID:3031656

A;Accession: A26460

A;Molecule type: DNA

A;Residues: 1-49 <DE>

A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:G182686; PIDN:AA5337

R;Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A;Title: Evolution of the fibronectin gene.

A;Reference number: A26284; MUID:86111901; PMID:3003095

A;Accession: A26284

A;Molecule type: DNA

A;Residues: 1447-1540 <OLD>

A;Cross-references: GB:M12549; NID:G182688

A;Note: the authors translated the codon TTC for residue 1494 as Glu

R;Paolella, G.; Henschcliffe, C.; Sebastio, G.; Baralle, F.E.

Nucleic Acids Res. 16, 3545-3557, 1988

A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B

A;Reference number: S00848; MUID:88233940; PMID:3375063

A;Accession: S03917

A;Molecule type: DNA

A;Residues: 1594-1767, 'V', 1769-1783 <PAO>

A;Cross-references: EMBL:X07719; NID:G31402

A;Note: the authors translated the codon AAC for residue 1631 as Asp

R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.

FEBS Lett. 207, 287-291, 1986

A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene:

A;Reference number: A24854; MUID:87030929; PMID:3770201

A;Accession: A24854

A;Molecule type: DNA

A;Residues: 1992-2147 <VIB>

A;Cross-references: GB:X04530; NID:G31436

R;Gutman, A.; Yamada, K.M.; Kornblihtt, A.

FEBS Lett. 207, 145-148, 1986

A;Title: Human fibronectin is synthesized as a pre-propolypeptide.

A;Reference number: A24476; MUID:87030890; PMID:3770189

A;Accession: A24476

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-14, 'Q', 16-38 <GUT>

R;Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.

EMBO J. 4, 1755-1759, 1985

A;Title: Primary structure of human fibronectin: differential splicing may generate at 1

notch protein - fr
thrombospondin 1 p
C-Delta-1 - chicke
Xotch protein - Af
notch-1 protein -
phospholipase-A(2)
notch protein homo
crumbs protein - f
hypothetical prote
probable laminin a
coagulation factor
laminin alpha-1 ch
insulin-like growt
cation-independent
cation-independent
transmembrane prot

A;Reference number: A91008; MUID:85284965; PMID:2992939
A;Accession: A91008
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 32-1344,1346-2080;2112-2386 <KOR>
R;Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptide
A;Reference number: A93529; MUID:84272258; PMID:6462919
A;Accession: A93529
A;Molecule type: mRNA
A;Residues: 973-2080;2112-2386 <KO2>
R;Cross-references: GB:X00739
F;Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
A;Reference number: A21011; MUID:83290929; PMID:6688418
A;Accession: A21011
A;Molecule type: mRNA
A;Residues: 1434-1537 <OL2>
R;Cross-references: GB:X00055; NID:g182680; PIDN:AAA52459.1; PID:g182683
R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
A;Reference number: A90495; MUID:85280409; PMID:2992573
A;Accession: A90495
A;Molecule type: mRNA
A;Residues: 1594-2386

R;Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697
R;Umezawa, K.; Kornblith, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A;Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A;Reference number: A22245; MUID:85231203; PMID:2989004
A;Accession: A22245
A;Molecule type: mRNA
A;Residues: 1948-2067 <UME>
R;Cross-references: GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:g182706
A;Accession: B22245
A;Molecule type: mRNA
A;Residues: 1975-1991;2017-2039 <UM2>
R;Cross-references: GB:M27590
R;Sekiguchi, K.; Kios, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A;Title: Human liver fibronectin complementary DNAs: identification of two different me
A;Reference number: I52394; MUID:87026578; PMID:3021206
A;Accession: I65273
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1978-1990,2016-2018, 'N', '2020-2081,2113-2127 <SEK>
R;Cross-references: GB:M4060; NID:g182701; PIDN:AAA52464.1; PID:g182704
R;Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectins
A;Reference number: A21165; MUID:83221567; PMID:6304699
A;Accession: A21165
A;Molecule type: mRNA
A;Residues: 2291-2386 <K03>
R;Cross-references: GB:X00799; NID:g182681; PIDN:AAA52460.1; PID:g182684
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A;Title: Primary structure of human plasma fibronectin.
A;Reference number: A92398; MUID:84032463; PMID:6630202
A;Accession: A92398
A;Molecule type: protein
A;Residues: 32-47, 'C', '49-51, 'S', '53-72, 'A', '74-290 <GAR1>
R;Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A;Title: Further characterization of the binding of fibronectin to gelatin reveals the F
A;Reference number: S34791; MUID:93312001; PMID:8323285
A;Accession: S34791
A;Molecule type: protein
A;Residues: 291-300;551-560 <GAR2>

R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A;Reference number: A60904; MUID:87019725; PMID:3532418
A;Accession: A60904
A;Molecule type: protein
A;Residues: 293-301 <GRI>
R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Sirl, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human p1
A;Reference number: A23901; MUID:86008277; PMID:3900070
A;Accession: A23901
A;Molecule type: protein
A;Residues: 616-677, 'Q', '679-703, 'PT' <CAL>
R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
A;Reference number: A92386; MUID:82265604; PMID:7050098
A;Accession: A92386
A;Molecule type: protein
A;Residues: 1441-1548 <PIE>
A;Note: residues 1524-1527 are responsible for the cell-binding activity
R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A;Reference number: A32517; MUID:87241275; PMID:3593230
A;Accession: A32517
A;Molecule type: protein
A;Residues: 1589-1630, 'T', '1722-2058 <GAR3>
R;Tressell, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A;Title: Human plasma fibronectin. Demonstration of structural differences between the A
A;Reference number: S14357; MUID:91190085; PMID:2012601
A;Accession: S14357
A;Molecule type: protein
A;Residues: 1614-1630, 'T', '1722-2081,2113-2244 <TRE>
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A;Reference number: A23891; MUID:85261459; PMID:4019516
A;Accession: A23891
A;Molecule type: protein
A;Residues: 2071-2080;2112-2356 <GAR4>
C;Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
action, and transformation.
C;Genetics:
A;Gene: GDB:FNI
A;Cross-references: GDB:I19135; OMIM:135600
A;Map position: 2q34-2q34
A;Inrons: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-31/Domain: propeptide #status predicted <PRO>
F;32-2386/Product: fibronectin #status experimental <MAT>
F;52-272/Domain: fibrin and heparin binding <FHB>
F;52-87/Domain: fibronectin type I repeat homology <1F1>
F;57-135/Domain: fibronectin type I repeat homology <1F2>
F;141-179/Domain: fibronectin type I repeat homology <1F3>
F;186-225/Domain: fibronectin type I repeat homology <1F4>
F;231-270/Domain: fibronectin type I repeat homology <1F5>
F;308-608/Domain: collagen binding <CBR>
F;308-342/Domain: fibronectin type I repeat homology <1F6>
F;360-401/Domain: fibronectin type II repeat homology <2F1>
F;420-461/Domain: fibronectin type II repeat homology <2F2>
F;470-508/Domain: fibronectin type I repeat homology <1F7>
F;518-555/Domain: fibronectin type I repeat homology <1F8>
F;561-589/Domain: fibronectin type I repeat homology <1F9>
F;609-692/Domain: fibronectin type III repeat homology <3FA>
F;616-706/Domain: heparin binding <HPB>
F;719-801/Domain: fibronectin type III repeat homology <3FB>
F;810-891/Domain: fibronectin type III repeat homology <3FC>

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 223.451 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-1

Perfect score: 3781

Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILRWRPVSPPRNLGY 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/PCT_NEW PUB.pdb.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW PUB.pdb.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pdb.*
4: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pdb.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB.pdb.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_NEW PUB.pdb.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pdb.*
8: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pdb.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_NEW PUB.pdb.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_NEW PUB.pdb.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_NEW PUB.pdb.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB.pdb.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_NEW PUB.pdb.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_NEW PUB.pdb.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_NEW PUB.pdb.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_NEW PUB.pdb.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_NEW PUB.pdb.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_NEW PUB.pdb.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_NEW PUB.pdb.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW PUB.pdb.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB.pdb.*
22: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3664	96.9	642	16 US-10-741-601-354	Sequence 354, App
2	3664	96.9	642	17 US-10-741-600-1066	Sequence 1066, App
3	3646.5	96.4	657	16 US-10-741-601-359	Sequence 359, App
4	3646.5	96.4	657	17 US-10-741-600-1072	Sequence 1072, App
5	3599.5	95.2	2220	15 US-10-236-392-4	Sequence 4, Appli
6	3599.5	95.2	2355	15 US-10-144-194A-104	Sequence 104, App
7	3599.5	95.2	2355	15 US-10-447-161-3	Sequence 3, Appli
8	3599.5	95.2	2355	16 US-10-734-564-94	Sequence 94, Appli
9	3599.5	95.2	2355	16 US-10-491-566-104	Sequence 104, App
10	3599.5	95.2	2355	17 US-10-852-335A-147	Sequence 147, App
11	3599.5	95.2	2355	18 US-10-287-436A-436	Sequence 436, App

12	3599.5	95.2	2355	18	US-10-287-436A-1137	Sequence 1137, App
13	3599.5	95.2	2355	20	US-11-040-130-28	Sequence 28, Appl
14	3599.5	95.2	2386	16	US-10-618-281-32	Sequence 32, Appl
15	3599.5	95.2	2474	18	US-10-450-763-52638	Sequence 52638, A
16	3595.5	95.1	2386	10	US-09-961-403-1	Sequence 1, Appli
17	3595.5	95.1	2386	16	US-10-788-792-206	Sequence 206, App
18	3595.5	95.1	2386	17	US-10-868-577A-59	Sequence 59, Appl
19	3595.5	95.1	2386	18	US-10-485-758-4	Sequence 4, Appli
20	3595.5	95.1	2386	18	US-10-485-758-9	Sequence 9, Appli
21	3592.5	95.0	984	16	US-10-741-601-356	Sequence 356, App
22	3592.5	95.0	984	17	US-10-741-600-1069	Sequence 1069, Ap
23	3592.5	95.0	2296	16	US-10-741-601-363	Sequence 363, App
24	3592.5	95.0	2296	17	US-10-741-600-1075	Sequence 1075, Ap
25	3592.5	95.0	2355	16	US-10-741-601-357	Sequence 357, App
26	3592.5	95.0	2355	16	US-10-741-601-366	Sequence 366, App
27	3592.5	95.0	2355	17	US-10-741-600-1067	Sequence 1067, Ap
28	3592.5	95.0	2355	17	US-10-741-600-1078	Sequence 1078, Ap
29	3592.5	95.0	2386	16	US-10-741-601-360	Sequence 360, App
30	3592.5	95.0	2386	17	US-10-741-600-1071	Sequence 1071, Ap
31	3448.5	91.2	2328	14	US-10-171-311-64	Sequence 64, Appl
32	3448.5	91.2	2328	15	US-10-236-031B-70	Sequence 70, Appl
33	3448.5	91.2	2328	15	US-10-374-979-98	Sequence 98, Appl
34	3448.5	91.2	2328	15	US-10-182-936A-98	Sequence 98, Appl
35	3448.5	91.2	2328	16	US-10-477-238A-677	Sequence 677, App
36	3448.5	91.2	2328	16	US-10-680-287A-677	Sequence 677, App
37	3448.5	91.2	2328	17	US-10-477-173-677	Sequence 235, App
38	3439.5	91.0	2355	15	US-10-360-101-235	Sequence 8, Appli
39	3424.5	90.6	2320	14	US-10-379-733-8	Sequence 2, Appli
40	3410.5	90.2	2320	15	US-10-236-392-2	Sequence 52634, A
41	3236.5	87.2	1173	18	US-10-450-763-52634	Sequence 52, Appl
42	2144	56.7	463	15	US-10-144-194A-52	Sequence 52, Appl
43	2144	56.7	463	16	US-10-491-566-52	Sequence 1, Appli
44	1859.5	49.2	343	9	US-09-934-706-1	Sequence 5, Appli
45	1859.5	49.2	400	9	US-09-934-706-5	

ALIGNMENTS

RESULT 1
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match	96.9%	Score	3664	DB	16	Length	642
Best Local Similarity	99.5%	Pred. No.	8.5e-289				
Matches	639	Conservative	0	Mismatches	3	Indels	0
Gaps	0						
Qy	19	MLRGGPGLLLAVQLCGTAVPSTGASKSKRQAOQMVQPSVAVSQSKPCYDNGKHQY	78				
Db	1	MLRGGPGLLLAVQLCGTAVPSTGASKSKRQAOQMVQPSVAVSQSKPCYDNGKHQY	60				
Qy	79	INQWERTYLGALVCTCYGSGRGNCEKPEABETCFDKYTGNTYRVGDTYERPKDSMI	138				
Db	61	INQWERTYLGALVCTCYGSGRGNCEKPEABETCFDKYTGNTYRVGDTYERPKDSMI	120				
Qy	139	WDCTCICAGRGRTSCTTANRCHGGQSKYKIGDTRRRPHETGGYMLCVCCLNGKGWTCCK	198				
Db	121	WDCTCICAGRGRTSCTTANRCHGGQSKYKIGDTRRRPHETGGYMLCVCCLNGKGWTCCK	180				

Qy 199 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNRCNDQDTRTSY 258
Db 181 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNRCNDQDTRTSY 240
Qy 259 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 318
Db 241 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 300
Qy 319 QPPPYGHCVTDSGVVYSVGMQWLKTQGNKQMLCTCLNGVSCQETA VTTQTYGGNSGEP 378
Db 301 QPPPYGHCVTDSGVVYSVGMQWLKTQGNKQMLCTCLNGVSCQETA VTTQTYGGNSGEP 360
Qy 379 VLPTYNDRDTSYTSNVEODQKYSFCTDHTVLVQTRGNSGALCHFPFLYNNHNYTDC 438
Db 361 VLPTYNDRDTSYTSNVEODQKYSFCTDHTVLVQTRGNSGALCHFPFLYNNHNYTDC 420
Qy 439 SEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDWDKQHDGMHMR 498
Db 421 SEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDWDKQHDGMHMR 480
Qy 499 CTCVNGRGWTCIAYSQLRDQCI VDDITVNVNDTFHKEHBEHMLNCTCFGQGRWKC 558
Db 481 CTCVNGRGWTCIAYSQLRDQCI VDDITVNVNDTFHKEHBEHMLNCTCFGQGRWKC 540
Qy 559 DPVQCODSETGYOIGDSWEKYVHGVRQCYCYGRGIGEMHCOPLQTYPSSSGPVEVF 618
Db 541 DPVQCODSETGYOIGDSWEKYVHGVRQCYCYGRGIGEMHCOPLQTYPSSSGPVEVF 600
Qy 619 ITETPSQNSHPIQWNAPOPSHISKYILRWRPVSIPPRNLGY 660
Db 601 ITETPSQNSHPIQWNAPOPSHISKYILRWRPVSIPPRNLGY 642

RESULT 2
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match 96.9%; Score 3664; DB 17; Length 642;
Best Local Similarity 99.5%; Pred. No. 8.5e-289;
Matches 639; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 19 MLRGPGLLLAVQCLGTAVPSTGASKSKRQAQMVQPSVAVSQSKPGCYDNGKHQY 78
Db 1 MLRGPGLLLAVQCLGTAVPSTGASKSKRQAQMVQPSVAVSQSKPGCYDNGKHQY 60
Qy 79 INQWERTYLGNAVCTCYGSGRGNCEKPEAETCFDKYTGNTYRVDYERPKDSMI 138
Db 61 INQWERTYLGNAVCTCYGSGRGNCEKPEAETCFDKYTGNTYRVDYERPKDSMI 120
Qy 139 WDCTCIGAGRISCTIANRCHGGQSYKIGDTRWRPHETGMYLCEVCLNGKGEWTC 198
Db 121 WDCTCIGAGRISCTIANRCHGGQSYKIGDTRWRPHETGMYLCEVCLNGKGEWTC 180
Qy 199 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNRCNDQDTRTSY 258
Db 181 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNRCNDQDTRTSY 240

Qy 259 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 318
Db 241 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 300
Qy 319 QPPPYGHCVTDSGVVYSVGMQWLKTQGNKQMLCTCLNGVSCQETA VTTQTYGGNSGEP 378
Db 301 QPPPYGHCVTDSGVVYSVGMQWLKTQGNKQMLCTCLNGVSCQETA VTTQTYGGNSGEP 360
Qy 379 VLPTYNDRDTSYTSNVEODQKYSFCTDHTVLVQTRGNSGALCHFPFLYNNHNYTDC 438
Db 361 VLPTYNDRDTSYTSNVEODQKYSFCTDHTVLVQTRGNSGALCHFPFLYNNHNYTDC 420
Qy 439 SEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDWDKQHDGMHMR 498
Db 421 SEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDWDKQHDGMHMR 480
Qy 499 CTCVNGRGWTCIAYSQLRDQCI VDDITVNVNDTFHKEHBEHMLNCTCFGQGRWKC 558
Db 481 CTCVNGRGWTCIAYSQLRDQCI VDDITVNVNDTFHKEHBEHMLNCTCFGQGRWKC 540
Qy 559 DPVQCODSETGYOIGDSWEKYVHGVRQCYCYGRGIGEMHCOPLQTYPSSSGPVEVF 618
Db 541 DPVQCODSETGYOIGDSWEKYVHGVRQCYCYGRGIGEMHCOPLQTYPSSSGPVEVF 600
Qy 619 ITETPSQNSHPIQWNAPOPSHISKYILRWRPVSIPPRNLGY 660
Db 601 ITETPSQNSHPIQWNAPOPSHISKYILRWRPVSIPPRNLGY 642

RESULT 3
US-10-741-601-359
; Sequence 359, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-359

Query Match 96.4%; Score 3646.5; DB 16; Length 657;
Best Local Similarity 97.3%; Pred. No. 2.3e-287;
Matches 639; Conservative 0; Mismatches 3; Indels 15; Gaps 1;
Qy 19 MLRGPGLLLAVQCLGTAVPSTGASKSKRQAQMVQPSVAVSQSKPGCYDNGKHQY 78
Db 1 MLRGPGLLLAVQCLGTAVPSTGASKSKRQAQMVQPSVAVSQSKPGCYDNGKHQY 60
Qy 79 INQWERTYLGNAVCTCYGSGRGNCEKPEAETCFDKYTGNTYRVDYERPKDSMI 138
Db 61 INQWERTYLGNAVCTCYGSGRGNCEKPEAETCFDKYTGNTYRVDYERPKDSMI 120
Qy 139 WDCTCIGAGRISCTIANRCHGGQSYKIGDTRWRPHETGMYLCEVCLNGKGEWTC 198
Db 121 WDCTCIGAGRISCTIANRCHGGQSYKIGDTRWRPHETGMYLCEVCLNGKGEWTC 180
Qy 199 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNRCNDQDTRTSY 258
Db 181 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNRCNDQDTRTSY 240
Qy 259 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 318
Db 241 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 300
Qy 319 QPPPYGHCVTDSGVVYSVGMQWLKTQGNKQMLCTCLNGVSCQETA VTTQTYGGNSGEP 378

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 69.8824 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-1

Perfect score: 3781

Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILWRPVSIPPRNLGY 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*

5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3595.5	95.1	2386	4	US-09-961-403-1
2	3579.5	94.7	2386	2	US-09-016-366A-12
3	3578.5	94.6	2446	2	US-08-551-356-2
4	3578.5	94.6	2446	5	PCT-US93-12687-2
5	3433.5	90.8	2327	6	5455158-1
6	3433.5	90.8	2327	6	5455158-1
7	3422.5	90.5	2324	1	US-08-283-857-1
8	3422.5	90.5	2324	5	PCT-US95-09819-1
9	3406.5	90.1	2231	1	US-08-153-799-16
10	371	9.8	65	2	US-08-982-597A-19
11	371	9.8	65	3	US-09-136-218-19
12	298	7.9	663	4	US-09-194-468A-30
13	297	7.9	631	3	US-08-448-489-17
14	297	7.9	631	4	US-09-689-730-17
15	297	7.9	660	3	US-08-704-711A-18
16	297	7.9	660	3	US-09-521-220-18
17	297	7.9	660	3	US-09-391-104-19
18	297	7.9	660	4	US-09-917-254-89
19	297	7.9	660	4	US-09-949-016-6512
20	297	7.9	660	4	US-09-949-016-7937
21	295	7.8	422	2	US-08-836-854-12
22	294.5	7.8	188	1	US-08-142-449B-14
23	294.5	7.8	474	2	US-08-836-854-9
24	292.5	7.7	429	4	US-09-194-468A-45
25	290	7.7	48	2	US-08-982-597A-22
26	290	7.7	48	3	US-09-136-218-22
27	285.5	7.6	707	3	US-08-704-711A-19
28					Sequence 1, Appl
29					Sequence 12, Appl
30					Sequence 2, Appl
31					Sequence 2, Appl
32					Patent No. 5455158
33					Patent No. 5455158
34					Sequence 1, Appl
35					Sequence 16, Appl
36					Sequence 19, Appl
37					Sequence 19, Appl
38					Sequence 30, Appl
39					Sequence 17, Appl
40					Sequence 17, Appl
41					Sequence 18, Appl
42					Sequence 19, Appl
43					Sequence 89, Appl
44					Sequence 6512, Ap
45					Sequence 7937, Ap
46					Sequence 12, Appl
47					Sequence 14, Appl
48					Sequence 9, Appl
49					Sequence 45, Appl
50					Sequence 22, Appl
51					Sequence 22, Appl
52					Sequence 19, Appl

Sequence 19, Appl
Sequence 20, Appl
Sequence 6575, Ap
Sequence 16, Appl
Sequence 16, Appl
Sequence 10629, A
Sequence 20, Appl
Sequence 20, Appl
Sequence 21, Appl
Sequence 21, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 18, Appl
Sequence 5, Appl
Sequence 17, Appl
Sequence 17, Appl
Sequence 168, App

ALIGNMENTS

RESULT 1

US-09-961-403-1

; Sequence 1, Application US/09961403

; Patent No. 6780594

; GENERAL INFORMATION:

; APPLICANT: HE-STUMPP, HOLGER

; APPLICANT: HAENDLER, BERNARD

; APPLICANT: KRAETZSCHMAR, JOERN

; APPLICANT: KREFT, BERTHOLT

; APPLICANT: WINTERHAGER, ELKE

; APPLICANT: REGIDOR, PEDRO

; APPLICANT: SCOTTI, SIMONE

; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS

; FILE REFERENCES: SCH-1789

; CURRENT APPLICATION NUMBER: US/09/961,403

; CURRENT FILING DATE: 2001-09-25

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2386

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-961-403-1

Query Match 95.1%; Score 3595.5; DB 4; Length 2386;

Best Local Similarity 97.2%; Pred. No. 7.2e-301;

Matches 629; Conservative 1; Mismatches 2; Indels 15; Gaps 1;

QY 19 MLRGPQPGLLLLAVQCLGTAVPSTGASKKROAQWVQPSVAVSQSKPGCYDNGKHQY 78

Db 1 MLRGPQPGLLLLAVQCLGTAVPSTGASKKROAQWVQPSVAVSQSKPGCYDNGKHQY 60

QY 79 INQWERTYLGALVCTCYGSGRGFNCESKPEAEETCFDKYTGNTYRVGDTYERPKDSMI 138

Db 61 INQWERTYLGALVCTCYGSGRGFNCESKPEAEETCFDKYTGNTYRVGDTYERPKDSMI 120

QY 139 WDCTCTGAGRGTSCTIANRCHEGGQSYKIGDTRWRPHETGGYMLCVCVCLNGKGWTCCK 198

Db 121 WDCTCTGAGRGTSCTIANRCHEGGQSYKIGDTRWRPHETGGYMLCVCVCLNGKGWTCCK 180

QY 199 PIAKCFDHAAGTSYVVGTEWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 258

Db 181 PIAKCFDHAAGTSYVVGTEWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240

QY 259 RIGDTSWKKDNRNLLQCLCTGNRGWEKWCERHTSVQTTSSGSGPFTDVRAAVYQPPHP 318

Db 241 RIGDTSWKKDNRNLLQCLCTGNRGWEKWCERHTSVQTTSSGSGPFTDVRAAVYQPPHP 300

QY 319 QPPPYGHCVTDSVYVSVGNQWMLCTGKNGVSCQRTAVTQTYGNSNGEPC 378

|||||

Db 301 QPPYGHCVTDGSGVYSGVMQMLKTQGNKQMLCTCLNGVSCQETAVTQTQTYGNSNGEPC 360
Qy 379 VLPFTYNDRT-----DSTTSNYEQDKYSFCTDHTVLVQTRGNSNGALC 423
Db 361 VLPFTYNGRTFYSCCTTEGRQDGHLMWCSTTSNYEQDKYSFCTDHTVLVQTRGNSNGALC 420
Qy 424 HFPELYNNHNTDCTSEGRDNDNMKCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 483
Db 421 HFPELYNNHNTDCTSEGRDNDNMKCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 480
Qy 484 GDQWDKQHDGMHMRCTCVNGRGWETCIAYSQLRDQCI VDDIITNNVNDTFFKHEEGHM 543
Db 481 GDQWDKQHDGMHMRCTCVNGRGWETCIAYSQLRDQCI VDDIITNNVNDTFFKHEEGHM 540
Qy 544 LNCTCFQGRGRWKCDPVDQCDSETGTFYIGDSWEKYVGVRYQCYCGRGIGEWHCQ 603
Db 541 LNCTCFQGRGRWKCDPVDQCDSETGTFYIGDSWEKYVGVRYQCYCGRGIGEWHCQ 600
Qy 604 PLQTPSSSGPVEVFTETPSQPNSHPIQWNAPOPSHISKYLWRP 650
Db 601 PLQTPSSSGPVEVFTETPSQPNSHPIQWNAPOPSHISKYLWRP 647

RESULT 2

US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2386 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-016-366A-12

Query Match 94.7%; Score 3579.5; DB 2; Length 2386;
Best Local Similarity 96.8%; Pred. No. 1.7e-299;
Matches 626; Conservative 1; Mismatches 5; Indels 15; Gaps 1;

Qy 19 MLRGPGLILLVQCLGTAVPSTGASKSRQQAQVQPOSQVAVSQSKPCYDNGKHQY 78
Db 1 MLRGPGLILLVQCLGTAVPSTGASKSRQQAQVQPOSQVAVSQSKPCYDNGKHQY 60
Qy 79 INQWERTYLGNAIVCTCYGSGRGFNCESKPEAEETCFDKYTGNTYRVGDTYERPKDSMI 138
Db 61 INQWERTYLGNAIVCTCYGSGRGFNCESKPEAEETCFDKYTGNTYRVGDTYERPKDSMI 120
Qy 139 WDCTCIGAGRGISCTIANRCHGGGOSYKIGDTRRRPHETGGYMLCEVCLGNGKGWTC 198
Db 121 WDCTCIGAGRGISCTIANRCHGGGOSYKIGDTRRRPHETGGYMLCEVCLGNGKGWTC 180
Qy 199 PIAEKCFDHAAGTSYVVGETWEKPYQGMVMDCTCLGEGSGRITCTSRNCNDQDTRTSY 258
Db 181 PIAEKCFDHAAGTSYVVGETWEKPYQGMVMDCTCLGEGSGRITCTSRNCNDQDTRTSY 240
Qy 259 RIGDTSKDNKRNGLQCICTGNGRGWKCERHTSVQTTSSGSGPFTDVRAAVYQPPHP 318
Db 241 RIGDTSKDNKRNGLQCICTGNGRGWKCERHTSVQTTSSGSGPFTDVRAAVYQPPHP 300
Qy 319 QPPYGHCVTDGSGVYSGVMQMLKTQGNKQMLCTCLNGVSCQETAVTQTQTYGNSNGEPC 378
Db 301 QPPYGHCVTDGSGVYSGVMQMLKTQGNKQMLCTCLNGVSCQETAVTQTQTYGNSNGEPC 360
Qy 379 VLPFTYNDRT-----DSTTSNYEQDKYSFCTDHTVLVQTRGNSNGALC 423
Db 361 VLPFTYNGRTFYSCCTTEGRQDGHLMWCSTTSNYEQDKYSFCTDHTVLVQTRGNSNGALC 420
Qy 424 HFPELYNNHNTDCTSEGRDNDNMKCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 483
Db 421 HFPELYNNHNTDCTSEGRDNDNMKCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 480
Qy 484 GDQWDKQHDGMHMRCTCVNGRGWETCIAYSQLRDQCI VDDIITNNVNDTFFKHEEGHM 543
Db 481 GDQWDKQHDGMHMRCTCVNGRGWETCIAYSQLRDQCI VDDIITNNVNDTFFKHEEGHM 540
Qy 544 LNCTCFQGRGRWKCDPVDQCDSETGTFYIGDSWEKYVGVRYQCYCGRGIGEWHCQ 603
Db 541 LNCTCFQGRGRWKCDPVDQCDSETGTFYIGDSWEKYVGVRYQCYCGRGIGEWHCQ 600
Qy 604 PLQTPSSSGPVEVFTETPSQPNSHPIQWNAPOPSHISKYLWRP 650
Db 601 PLQTPSSSGPVEVFTETPSQPNSHPIQWNAPOPSHISKYLWRP 647

RESULT 3
US-08-551-356-2
; Sequence 2, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551.356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 244.157 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651d-1

Perfect score: 3781

Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILRWKRVSIIPRNLGY 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3781	100.0	660	2	AAY28901 Human mig
2	3693.5	97.7	2182	8	ADR66462 Human pro
3	3693.5	97.7	2182	8	ADR66120 Human pro
4	3680.5	97.3	720	2	AAY28914 Fibronec
5	3680	97.3	642	8	ADR67316 Human bla
6	3680	97.3	642	8	ADR67316 Human bla
7	3680	97.3	642	8	ADR97658 Human fib
8	3664	96.9	642	8	ADR97658 Human fib
9	3658	96.7	2447	4	AAW40434 Human pol
10	3658	96.7	2447	4	AAW40435 Human pol
11	3658	96.7	2447	4	AAW40433 Human pol
12	3658	96.7	2447	4	AAW40432 Human pol
13	3646.5	96.4	657	8	ADR39409 Human pro
14	3599.5	95.2	2220	6	ABO01289 Human cel
15	3599.5	95.2	2266	6	ABO01289 Human cel
16	3599.5	95.2	2355	6	ABR58335 NM_00202
17	3599.5	95.2	2355	7	ADP65196 Human fib
18	3599.5	95.2	2355	8	ADG89560 Human fib
19	3599.5	95.2	2355	8	ADO55175 Protein #
20	3599.5	95.2	2355	8	ADQ26085 Fibronec
21	3599.5	95.2	2355	8	ADQ29668 Human col
22	3599.5	95.2	2355	8	ADR67315 Human bla
23	3599.5	95.2	2386	8	ADO55174 Protein #
24	3599.5	95.2	2474	4	ABG22279 Novel hum
25	3599.5	95.1	2386	5	AAO17353 Human fib

26	3595.5	95.1	2386	6	ABR81866 Human fib
27	3595.5	95.1	2386	7	ADD18770 Human dis
28	3595.5	95.1	2386	7	ADG63324 Human pro
29	3595.5	95.1	2386	8	ADR99200 Fibronec
30	3593.5	95.0	2446	3	AB50377 Human fib
31	3592.5	95.0	984	8	ADQ39406 Human myo
32	3592.5	95.0	2265	4	AAW38647 Human pol
33	3592.5	95.0	2296	8	ADQ39412 Human myo
34	3592.5	95.0	2330	4	AAW38646 Human pol
35	3592.5	95.0	2355	4	AAW38649 Human pol
36	3592.5	95.0	2355	8	ADQ39415 Human myo
37	3592.5	95.0	2355	8	ADQ39404 Human myo
38	3592.5	95.0	2386	4	AAW38648 Human pol
39	3592.5	95.0	2386	8	ADQ39408 Human myo
40	3586.5	94.9	2477	2	AAW99595 Human fib
41	3584.5	94.8	2220	8	ADN95950 Human NOV
42	3579.5	94.7	2386	2	AAW63171 Amino aci
43	3578.5	94.6	2446	2	AAW60021 Fibrinoge
44	3546	93.8	2352	8	ADK00410 Fibronec
45	3448.5	91.2	2328	4	AAG68182 Fibronec

ALIGNMENTS

RESULT 1

AAV28901

ID AAV28901 standard; protein; 660 AA.

XX AAV28901;

XX 21-SEP-1999 (first entry)

XX Human migration stimulating factor (MSF) 1-alpha protein.

XX Migration stimulatory factor; MSF; cell migration; modulation; human;

XX wound healing; scarring; MSFI-alpha.

XX Homo sapiens.

XX WO9931233-A1.

XX 24-JUN-1999.

XX 15-DEC-1998; 98WO-GB003766.

XX 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI; 1999-430039/36.

XX N-PSDB; AAX81299.

Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

Claim 1; Page 53; 86pp; English.

The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. The present sequence represents the human MSFI-alpha protein

XX Sequence 660 AA;

Query Match 100.0%; Score 3781; DB 2; Length 660;

Best Local Similarity 100.0%; Pred. No. 7.1e-242; Indels 0; Gaps 0;

Matches 660; Conservative 0; Mismatches 0;

```
QY 1 NLVATCLPVRASLPRLNMLRPGPGLLLAVQCLGTAVPSTGASKSKRQAOQVQPOSP 60
Db 1 NLVATCLPVRASLPRLNMLRPGPGLLLAVQCLGTAVPSTGASKSKRQAOQVQPOSP 60
QY 61 VAVSQSKPGCYDNGKHQYQINQWERTYLGNALVCTCYGSGRGFNCSKPEAEETCFDKYT 120
Db 61 VAVSQSKPGCYDNGKHQYQINQWERTYLGNALVCTCYGSGRGFNCSKPEAEETCFDKYT 120
QY 121 GNTYRVGDTYRPPKDSMIWDCCTCAGRGRISCTIANRCHGGQSYKIGDTRWRPHETGG 180
Db 121 GNTYRVGDTYRPPKDSMIWDCCTCAGRGRISCTIANRCHGGQSYKIGDTRWRPHETGG 180
QY 181 YMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWMDCTCLGEGSGR 240
Db 181 YMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWMDCTCLGEGSGR 240
QY 241 ITCTSRNRNCNDQDTRTSYRIGDTSKDNKGNLLQICITGNRGEWKCEHRTSVQTTSSG 300
Db 241 ITCTSRNRNCNDQDTRTSYRIGDTSKDNKGNLLQICITGNRGEWKCEHRTSVQTTSSG 300
QY 301 SGPFTRVRAAVYQPPHPQPPYGHCVTDGVSYSVGNQWLKTQGNKQMLCTCLNGVSC 360
Db 301 SGPFTRVRAAVYQPPHPQPPYGHCVTDGVSYSVGNQWLKTQGNKQMLCTCLNGVSC 360
QY 361 QETAVTQYVGNSEGPCVLPFTYNDRTDSTTSNVEQDKYSFCTDHTVLVQTRGNSNG 420
Db 361 QETAVTQYVGNSEGPCVLPFTYNDRTDSTTSNVEQDKYSFCTDHTVLVQTRGNSNG 420
QY 421 ALCHFPFLYNNHNTDCTSEGRDNMKWCGTTQNYDADQKFGFCPMAAHEBICITNEGVM 480
Db 421 ALCHFPFLYNNHNTDCTSEGRDNMKWCGTTQNYDADQKFGFCPMAAHEBICITNEGVM 480
QY 481 YRIGQWDKQHDGMHMRCTCVGNRGEWTCIAYSQLRDQICVDITVNDTFRKHEE 540
Db 481 YRIGQWDKQHDGMHMRCTCVGNRGEWTCIAYSQLRDQICVDITVNDTFRKHEE 540
QY 541 GHMLNCTCFGQGRGWKCDPVQDQDSETGTYQIGDSWEKYVHGVRQCYCYGRGIGEW 600
Db 541 GHMLNCTCFGQGRGWKCDPVQDQDSETGTYQIGDSWEKYVHGVRQCYCYGRGIGEW 600
QY 601 HCQPLQTYPSSSGPVEVITETPSQPNHPIQWNAPOPSHISKYILRWRPVSIPRNLGY 660
Db 601 HCQPLQTYPSSSGPVEVITETPSQPNHPIQWNAPOPSHISKYILRWRPVSIPRNLGY 660

RESULT 2
ADR66462
ID ADR66462 standard; protein; 2182 AA.
XX
AC ADR66462;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human prostatic carcinoma derived protein SEQ ID 316 #2.
KW human; cytostatic; diagnosis; prostatic cancer;
KW differential expression analysis.
XX
OS Homo sapiens.
XX
PN WO2004076614-A2.
XX
PD 10-SEP-2004.
XX
PF 22-FEB-2004; 2004WO-DE000433.
XX
PR 27-FEB-2003; 2003DE-01009985.
PR 14-MAY-2003; 2003DE-01022134.
XX
(HINZ/) HINZMANN B.
PA (DAHL/) DAHL E.
PA (ROSE/) ROSENTHAL A.
PA (HERM/) HERMANN K.
```

(PILA/) PILARSKY C.

Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
Schmitt A, Beckmann G, Bruemendorf T, Kimmann H, Roepcke S;
Xinzhong L, Staub E;
WPI; 2004-653386/63.

New nucleic acids, and encoded proteins, from prostatic cancer tissue,
useful for diagnosis, treatment and in screening for specific binding
agents.

Claim 2; Page 1298; 1607pp; German.

This invention describes novel cytostatic polynucleotide and polypeptide
sequences which can be used in a method for diagnosing prostatic cancer
or the risk of developing prostatic cancer. Diagnosis is based on
determining over transcription or over expression of the sequences in
prostatic tissue. Screening for inhibitors of the sequences or detection
substances involves a binding assay, any compounds that bind are
selected, optionally after deconvolution of mixtures. Detection of a
predetermined minimum level of the reporter indicates the presence of
tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
short-interfering RNA or ribozymes; an organic molecule of molecular
weight below 5000, preferably 300, that binds to the polypeptide; an
aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
polypeptide, preferably humanised or human; an anti-idiotypic, non-human
(monoclonal) antibody directed against Ab or any of the above derivatised
with a reporter group, cell toxin, immunostimulatory molecules and/or
radioisotope. The polynucleotides are identified in human prostatic
cancer by differential expression analysis, using DNA microarrays,
between normal and tumorous tissues, with (over)expression being detected
by quantitative PCR. Analysis of prostatic cancer samples showed that
CD24 was upregulated in many of them. Sections of tissue, isolated from
prostatic cancer patients, or subjects at risk, were incubated
sequentially with anti-human CD4 murine monoclonal antibodies;
biotinylated second antibody; streptavidin-conjugated horseradish
peroxidase and then diaminobenzidine as colour former (brown). The
samples were counterstained with hemalum (blue). Malignant cells stained
strongly but non-malignant cells only weakly. In 15 of 63 samples of
adenocarcinoma, membrane and cytoplasmic staining was very strong, and
lymph node metastases were also stained. ADR65805-ADR6954 represent the
polynucleotide and polypeptide sequences used in the method of the
invention.

SQ Sequence 2182 AA;

Query Match 97.7%; Score 3693.5; DB 8; Length 2182;
Best Local Similarity 97.4%; Pred. No. 1.7e-235;
Matches 648; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 NLVATCLPVRASLPRLNMLRPGPGLLLAVQCLGTAVPSTGASKSKRQAOQVQPOSP 60
Db 71 NLVATCLPVRASLPRLNMLRPGPGLLLAVQCLGTAVPSTGASKSKRQAOQVQPOSP 130
QY 61 VAVSQSKPGCYDNGKHQYQINQWERTYLGNALVCTCYGSGRGFNCSKPEAEETCFDKYT 120
Db 131 VAVSQSKPGCYDNGKHQYQINQWERTYLGNALVCTCYGSGRGFNCSKPEAEETCFDKYT 190
QY 121 GNTYRVGDTYRPPKDSMIWDCCTCAGRGRISCTIANRCHGGQSYKIGDTRWRPHETGG 180
Db 191 GNTYRVGDTYRPPKDSMIWDCCTCAGRGRISCTIANRCHGGQSYKIGDTRWRPHETGG 250
QY 181 YMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWMDCTCLGEGSGR 240
Db 251 YMLECVCLNGKGWETCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWMDCTCLGEGSGR 310
QY 241 ITCTSRNRNCNDQDTRTSYRIGDTSKDNKGNLLQICITGNRGEWKCEHRTSVQTTSSG 300
Db 311 ITCTSRNRNCNDQDTRTSYRIGDTSKDNKGNLLQICITGNRGEWKCEHRTSVQTTSSG 370
QY 301 SGPFTRVRAAVYQPPHPQPPYGHCVTDGVSYSVGNQWLKTQGNKQMLCTCLNGVSC 360

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 13726.3 Seconds
(without alignments)
1830.243 Million cell updates/sec

Title: US-09-581-651D-1

Perfect score: 3781
Sequence: 1 NLVATCPVRASLPHRLML.....ISKYILWRPVSIPPRNLGY 660

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool_b/US09581651/runat.07112005.092223.28789/app_query.fasta.1.1834
-DB=EST -QFMT=fascap -SUPPIX=p2n.rst -MINWATCH=0.1 -LOOPC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CIGN 1.1 8180 @runat.07112005.092223.28789 -ICPU=3
-NO MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3685.5	97.5	7885	3 CR749281	CR749281 Homo sapi
2	3682.5	97.4	7777	3 CR749316	CR749316 Homo sapi
3	3661.5	96.8	8121	3 CR749317	CR749317 Homo sapi
4	3658.5	96.8	7501	3 BC078656	BC078656 Homo sapi
5	3419.5	90.4	8315	3 AK090135	AK090135 Mus muscu
6	3308	87.5	8329	3 AK090130	AK090130 Mus muscu
7	1573.5	41.6	943	5 BX391752	BX391752 BX391752
8	1564	41.4	1044	5 BX398837	BX398837 BX398837
9	1526	40.4	2107	3 AK054456	AK054456 Mus muscu

10	1490.5	39.4	836	7	CN538822	UI-M-HS0-
11	1480	39.1	1052	5	BX380582	BX380582
12	1478	39.1	816	1	AU141008	AU141008
13	1464	38.7	859	7	CF616056	AGENCOURT
14	1460	38.6	766	7	CN419594	170004345
15	1451.5	38.4	861	1	AU140971	AU140971
16	1450	38.3	1026	5	BX386270	BX386270
17	1419.5	37.5	781	7	CN534124	UI-M-HOO-
18	1415	37.4	765	4	BM715855	UI-E-EJO-
19	1402.5	37.1	1050	5	BX417945	BX417945
20	1400	37.0	808	7	CK638401	UI-M-HOO-
21	1390	36.8	837	1	AU140973	AU140973
22	1386	36.7	731	1	AU140526	AU140526
23	1381	36.5	737	1	AU140993	AU140993
24	1369	36.2	959	5	BU109952	603127860
25	1356	35.9	739	1	AL706215	DKFZp686J
26	1351	35.7	695	7	CN419477	170004247
27	1349	35.7	730	1	AU140834	AU140834
28	1340	35.4	772	7	CN530412	UI-M-HOO-
29	1329.5	35.2	781	7	CN161442	950735 MA
30	1323	35.0	699	1	AU140889	AU140889
31	1308	34.6	723	7	CN419622	170005319
32	1304.5	34.5	780	7	CN162843	952271 MA
33	1303	34.5	738	6	CD613781	55110394J
34	1302.5	34.4	748	1	AU140948	AU140948
35	1298	34.3	667	1	AU140910	AU140910
36	1295.5	34.3	796	1	AU140814	AU140814
37	1291.5	34.2	748	1	AU140991	AU140991
38	1289	34.1	676	1	AL603362	DKFZp686C
39	1284	34.0	705	1	AU140789	AU140789
40	1283.5	33.9	758	1	AU140556	AU140556
41	1282.5	33.9	941	5	BQ928772	AGENCOURT
42	1281	33.9	734	1	AU140802	AU140802
43	1279.5	33.8	851	1	AU141024	AU141024
44	1271	33.5	673	1	AU140450	AU140450
45	1245.5	32.9	759	1	AU140735	AU140735

ALIGNMENTS

RESULT 1
CR749281
LOCUS CR749281 Homo sapiens mRNA; cDNA DKFZp686F10164 (from clone DKFZp686F10164).
DEFINITION
ACCESSION CR749281
VERSION CR749281.1 GI:51476291
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 7885)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Pobo, G., Han, M. and Wiemann, S.
The German cDNA Consortium
Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by BPFZ (Biomedical Research Center at the Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686F10164) is available at the R2PD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact R2PD for ordering:
http://www.r2pd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686F10164
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
1..7885
/organism="Homo sapiens"
/mol_type="mRNA"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 3388.86 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-1
Perfect score: 3781
Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILWRPVSIPPRNLGY 660

Scoring table: BLOSUM62
GLOPOP 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1834
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPFCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl6sum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -WMAP=US09581651 @CGN 1 1 1041 @runat_07112005_092226_28867
-NCPU=6 -ICPU=3 -NO_MMAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DSLOP=6 -DEUEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3765	99.6	2127	22	US-10-956-157-4288
2	3759	99.4	2443	20	US-10-741-601-70
3	3759	99.4	2443	22	US-10-741-600-238
4	3758	99.4	2127	17	US-10-210-120-49
5	3758	99.4	2127	24	US-10-909-035-49
6	3741.5	99.0	2488	20	US-10-741-601-75
7	3741.5	99.0	2488	22	US-10-741-600-244
8	3693.5	97.7	8027	18	US-10-447-161-8
9	3693.5	97.7	8027	20	US-10-734-564-27
10	3693.5	97.7	8027	22	US-10-852-335A-53
11	3693.5	97.7	8027	24	US-10-287-436A-81
12	3693.5	97.7	8815	22	US-10-868-577A-62
13	3693.5	97.7	8815	20	US-10-868-549-21
14	3687.5	97.5	6510	22	US-10-741-601-72
15	3687.5	97.5	6510	20	US-10-741-600-241
16	3687.5	97.5	7823	20	US-10-741-601-77
17	3687.5	97.5	7823	22	US-10-741-600-245
18	3687.5	97.5	7848	20	US-10-741-601-78
19	3687.5	97.5	7848	22	US-10-741-600-246
20	3687.5	97.5	7935	20	US-10-741-601-74
21	3687.5	97.5	7935	22	US-10-741-600-240
22	3687.5	97.5	7959	20	US-10-741-601-81
23	3687.5	97.5	7959	22	US-10-741-600-249
24	3687.5	97.5	8013	20	US-10-741-601-71
25	3687.5	97.5	8013	22	US-10-741-600-242
26	3687.5	97.5	8155	20	US-10-741-601-79
27	3687.5	97.5	8155	22	US-10-741-600-247
28	3687.5	97.5	8226	20	US-10-741-601-69
29	3687.5	97.5	8226	22	US-10-741-600-237
30	3687.5	97.5	8278	20	US-10-741-601-82
31	3687.5	97.5	8278	22	US-10-741-600-250
32	3687.5	97.5	8332	20	US-10-741-601-73
33	3687.5	97.5	8332	22	US-10-741-600-239
34	3687.5	97.5	8371	20	US-10-741-601-76
35	3687.5	97.5	8371	22	US-10-741-600-243
36	3686.5	97.5	7795	16	US-10-084-817-2
37	3686.5	97.5	7867	14	US-10-098-841-6
38	3686.5	97.5	8044	17	US-10-240-965-121
39	3686.5	97.5	8044	24	US-10-765-700-135
40	3686.5	97.5	8062	14	US-10-098-841-5
41	3686.5	97.5	8137	14	US-10-098-841-8
42	3686.5	97.5	8230	14	US-10-098-841-7
43	3605.5	95.4	7361	19	US-10-236-392-3
44	3547.5	93.8	8216	24	US-10-450-763-22270
45	3527	93.3	4295	18	US-10-144-194A-51

ALIGNMENTS

RESULT 1
US-10-956-157-4288
; Sequence 4288, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCES: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956.157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: Patent version 3.2

Db 1803 CATTGCCAACCTTTACAGACCTATCCAGGCTCAAGTGGTCTCTCGAAGTATTTATCACT 1862
Qy 621 GluThrProSerGlnProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHis 640
Db 1863 GAGACTCCGAGTCAAGCCACTCCACCCCATCCAGTGGATGACACACAGCCATCTCAC 1922
Qy 641 IleSerLysTrpIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 660
Db 1923 ATTTCCAAGTACATCTCAGGTGAGACCTGTGAGTATCCACCCAGAAACCTTGGATAC 1982

RESULT 2

ADR67201

ID ADR67201 standard; DNA; 2358 BP.

XX AC ADR67201;

DT 02-DEC-2004 (first entry)

XX Human bladder cancer associated nucleotide sequence.

XX KW bladder cancer tissue; bladder cancer; cytostatic; gene; ds.

XX OS Homo sapiens.

XX PN W02004076613-A2.

XX PD 10-SEP-2004.

XX PF 24-FEB-2004; 2004WO-DE000364.

XX PR 26-FEB-2003; 2003DE-01009729.

XX PA (HERR/) HERR A.

XX PA (HINZ/) HINZMANN B.

XX PA (DAHL/) DAHL E.

XX PA (STAU/) STAU E.

XX PA (PILA/) PILARSKY C.

XX PA (SPEC/) SPECHT T.

XX PI Herr A, Hinzmann B, Dahl E, Staub E, Pilarsky C, Specht T;

XX WPI; 2004-653385/63.

DR XX New nucleic acids, and encoded proteins, from bladder cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
PT agents.

PS Claim 1; Fig 3; 112pp; German.

XX The present invention describes nucleic acids (I) associated with bladder
CC cancer tissue. Also described: (1) peptides and proteins (II) containing
CC an amino acid sequence encoded by (I); (2) a method for diagnosing
CC bladder cancer (BC), or monitoring its progression, that uses (I), (II)
CC or agents (Z) that inhibit, or bind to, (I) or (II), labelled with a
CC reporter; and (3) a method for treating BC that uses (I), (II) or (Z).
CC (I) and (II) have cytostatic activity. (I) and (II) can be used to detect
CC and monitor progression of) bladder cancer (BC), or the risk of
CC developing it; to screen for specific binding agents (Z), and to treat
CC BC. (Z) are also useful as diagnostic and therapeutic agents. The present
CC sequence represents a human nucleotide sequence associated with bladder
CC cancer, which is used in the exemplification of the present invention.

SQ Sequence 2358 BP; 611 A; 587 C; 639 G; 521 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.05e-245 Length: 2358
Score: 3774.00 Matches: 659
Percent Similarity: 99.85% Conservative: 0
Best Local Similarity: 99.85% Mismatches: 1
Query Match: 99.81% Indels: 0
Gaps: 13

US-09-581-651D-1 (1-660) x ADR67201 (1-2358)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 214 AACTTGGTGGCAACTTGCCTCCCGGTGGCGGTCTCTCCCCACCGCTCTCAACATGCTT 273
Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
Db 274 AGGGGTCCGGGGCCGGGCTGCTGCTGCTGGCGCTCCAGTGCCTGGGGACAGCGGTGCC 333
Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro 60
Db 334 TCCACGGGAGCCTCGAAGACCAAGAGCAGCGCTCAGCAATGGTTTCAGCCCGAGTCCCG 393
Qy 61 ValAlaValSerGlnSerLysProGlyCysArgAspAsnGlyLysHisIleThrGlnIleAsn 80
Db 394 GTGGCTGTCACTCAAGCAAGCCGGTTGTTATGACAAATGGAACAACTATCAGATAAT 453
Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysThrGlySer 100
Db 454 CAACAGTGGAGCGGACCTACCTAGGCAATGCGTGGTGGTGTGTTGTTATGGAGGAAGC 513
Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysThr 120
Db 514 CGAGGTTTTAACTCGGAGAGTAACCTGAAGCTGAAGAGACTTGTCTTTGACAGTACACT 573
Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
Db 574 GGGAAACACTTACCAGAGTGGGTGACACTTATGAGCGTCTTAAGAGACTCCATGATCTGGAC 633
Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
Db 634 TGTACCTGCATCGGGCTGGCGGAGGAGAAATAGCTGTACCATCGCAACCGCTGCCAT 693
Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180
Db 694 GAAGGGGTCACTCCCTACAAGATTGGTGACACTGGAGGAGACCACTAGACTGTGTGT 753
Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysProIle 200
Db 754 TACATGTTAGAGTGTGTGTCTTGGTAAATGGAAAGAGAGATGGACCTGCAAGCCCAT 813
Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
Db 814 GCTGAGAAGTGTGTTGATCATGCTGCTGGAGCTTCTTATGTTGGAGAACATGGGAG 873
Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db 874 AAGCCCTACCAAGCTGGATGATGGTAGATTGTACTTGGCTGGAGAAAGGAGCGGACGC 933
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgile 260
Db 934 ATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTATAGAAT 993
Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 994 GGAGACACCTCGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC 1053
Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1054 AAGCGCGGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGACAGACCACTCAGCGGA 1113
Qy 301 SerGlyProPheThrAspValArgAlaValTyrGlnProGlnProHisProGlnPro 320
Db 1114 TCTGGCCCTTTCACCGATGTTGTCGAGCTGTTTACCACCGCAGCCTCACCCCGCCCT 1173
Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 140
Db 1174 CTTCCCTATGGCCACTGTGTGCACAGACAGTGGTGTGTGTCTACTCTGTGGGATGCAGTGG 1233
Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1234 CTGAAGACACCAAGAAATAAGCAAAATGCTTTGACGCTGCTGGCAACCGAGTGCAGTGC 1293


```
Qy 361 GlnGluThrAlaValThrGlnThrTyrclyGlyAsnSerAsnGlyGluProCysValLeu 380
Db 1294 CAAGAGACAGCTGTAACCCAGACTTACGGTGGCACTCAAAATGGAGAGCCCATGTCTTTA 1353
Qy 381 ProPheThrTyrclyAsnAspAspThrThrSerAsnTyrclyGluGlnAspGlnlys 400
Db 1354 CCATTACCTCAATGGCAGGACGACAGCACTTCGAATTTATGACGAGCACAGAA 1413
Qy 401 TyrSerPheCysThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGly 420
Db 1414 TACTCTTCTTCAGACACACACTGTTTGTGTTCAGACTCGAGGAGGAAATTCCAATGGT 1473
Qy 421 AlaLeuCysHisPheProPheLeuTyrclyAsnHisAsnTyrclyThrAspCysThrSerGlu 440
Db 1474 GCCTTGTGCCACTTCCCTTCTTATACAAACACCAATTAACACTGATGTCACCTTCTGAG 1533
Qy 441 GlyArgArgAspAsnMetLysTrpCysGlyThrThrGlnAsnTyrclyAspAlaAspGlnlys 460
Db 1534 GGCAGAGAGACACATGAGTGTGTGGGACACACAGAACTATGATGCCGACACAGAG 1593
Qy 461 PheGlyPheCysProMetAlaAlaHisGluGluLeuLeuCysThrThrAsnGluGlyValMet 480
Db 1594 TTTGGGTTCTGCCCTCCATGCTGCCACGAGGAAATCTGCACAACTGAAGGGGTCTATG 1653
Qy 481 TyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThr 500
Db 1654 TACCGCATTTGAGATCAGTGGGATAGCAGCATGACATGGGTTCATGATGAGTGCACG 1713
Qy 501 CysValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrclySerGlnLeuArgAspGln 520
Db 1714 TGTGTGGGAATGGTGTGGGGAATGGACATGCTACTGCTACTCGACCTTCGAGATCAG 1773
Qy 521 CysIleValAspAspIleThrTyrclyAsnValAsnAspThrPheHisLysArgHisGluGlu 540
Db 1774 TGCATTTGTGATGACATCATTACAAATGTGAACGACACATTCACCAAGCGTCATGAAGAG 1833
Qy 541 GlyHisMetLeuAsnCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspPro 560
Db 1834 GGGCAGATGCTGAACATGTACATGCTTCGGTCAGGGTCGGGGCAGGTGGAAGTGTGATCCC 1893
Qy 561 ValAspGlnCysGlnAspSerGluThrGlyThrPheTyrclyGlnIleGlyAspSerTrpGlu 580
Db 1894 GTCGACCAATGCCAGGATTCAGAGCTGGGAGCTTTTATCAAAATGGAGATTCATGGGAG 1953
Qy 581 LysTyrclyValHisGlyValArgTyrclyGlnCysTyrclyArgGlyIleGlyGluTrp 600
Db 1954 AAGTATGTGTCATGGTGTGAGTACATGTCAGTGTCTACTGCTATGGCCGTGGCATTTGGGAGTGG 2013
Qy 601 HisCysGlnProLeuGlnThrTyrclyProSerSerGlyProValGluValPheIleThr 620
Db 2014 CATTCGCCAACCTTTACAGACCTTATCCAAAGCTCAAGTGGTCTCTGTCGAAGTATTTATCACT 2073
Qy 621 GluThrProSerGlnProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHis 640
Db 2074 GAGACTCCGAGTACGCCCACTCCACCCCATCCAGTGGATGACACACAGCCATCTCAC 2133
Qy 641 IleSerLysTyrclyIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyrcly 660
Db 2134 ATTTCCAAAGTACATTTCTCAGTGGGAGACCTGTGTGAGTATCCCAACCCAGGAAACCTTGGATAC 2193
RESULT 3
ADQ38575
ID ADQ38575 standard; DNA; 2443 BP.
XX AC ADQ38575;
XX XX
DT 18-NOV-2004 (first entry)
XX DE Human SNP containing myocardial infarction-associated gene, SEQ ID 238.
XX DE Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX KW cardiant; gene therapy; human; gene; ds.
KW KW
```

```
XX Homo sapiens.
OS
FN WO2004058052-A2.
XX
PD 15-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US040978.
XX
PR 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Cargill M, Devlin JJ, Iakoubova O;
XX WPI; 2004-533949/51.
DR P-PSDB; ADQ39403.
XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 7; SEQ ID NO 238; 145pp; English.
XX
CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX
SQ Sequence 2443 BP; 599 A; 628 C; 682 G; 532 T; 0 U; 2 Other;
```

```
Alignment Scores:
Pred. No.: 1,13e-244 Length: 2443
Score: 3759.00 Matches: 657
Percent Similarity: 99.55% Conservativity: 0
Best Local Similarity: 99.55% Mismatches: 3
Query Match: 99.42% Indels: 0
DB: 13 Gaps: 0
```

US-09-581-651D-1 (1-660) x ADQ38575 (1-2443)

```
Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 319 AACTTGTGGCACTTGGCTCCCGTGGGGCGTCTCTCCCCACCGCTCAACATGCTT 378
Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
|||||
```



```
PD 13-FEB-2003.
XX
XX 02-AUG-2002; 2002WO-US024567.
PF
XX 02-AUG-2001; 2001US-0309581P.
PR 15-NOV-2001; 2001US-0334468P.
PR 01-AUG-2002; 2002US-00210120.
XX
XX (UNMI ) UNIV MICHIGAN.
PA
XX
XX Rubin MA, Chinnaiyan AM, Sreekumar A;
XX WPI; 2003-278396/27.
XX
XX Characterizing prostate tissue comprises providing a prostate tissue
PT sample from a subject and detecting the presence or absence of expression
PT of hepsin, pim-1 or EZH2.
XX
XX Disclosure; SEQ ID NO 49; 297pp; English.
XX
XX This invention relates to a novel method of characterising prostate
CC tissue in a subject and to compositions and methods for cancer
CC diagnostics, including cancer markers, in particular prostate cancer.
CC Prostate cancer (PCA) is a leading cause of male cancer-related death.
CC Additional serum and tissue biomarkers would aid diagnosis. The invention
CC may provide means of producing compounds with a cytostatic activity or
CC allow the development of gene therapy. The methods of the invention
CC useful for characterising prostate tissue in a subject, screening
CC compounds, characterising inconclusive prostate biopsy tissue in a
CC subject, detecting AMACR (alpha-methylacyl-coenzyme A racemase)
CC expression in a bodily fluid, characterising tissue in a subject,
CC diagnosing cancer in a subject and inhibiting the growth of cells. The
CC present sequence is a DNA sequence which is preferably utilised in the
CC method of the invention.
XX
XX
XX SQ Sequence 2127 BP; 552 A; 511 C; 576 G; 488 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1,126-244 Length: 2127
XX Score: 3758.00 Matches: 657
XX Percent Similarity: 99.55% Conservative: 0
XX Best Local Similarity: 99.55% Mismatches: 3
XX Query Match: 99.39% Indels: 0
XX DB: 10 Gaps: 0
XX
XX US-09-581-651D-1 (1-660) x ADD18477 (1-2127)
XX
XX 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
XX
XX 3 AACTTGGTGGCAACTTGGCTCCCGTGGCGGGTCTCTCCGCCACCGTCTCAACATGCTT 62
XX
XX 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
XX
XX 63 AGGGGTCCGGGGCCGGGTGCTGCTGGTGGCCGCTCTGCTGGGACAGCGGTGCC 122
XX
XX 41 SerThrGlyAlaSerIysSerIysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
XX
XX 123 TCCAGGGAGCCTCGAAGAGCAGAGGAGGCTAGCAAAATGGTTACGCCCCAGTCCCG 182
XX
XX 61 ValAlaValSerGlnSerIysProGlyCysThrAspAsnGlyIysHisIleThrGlnIleAsn 80
XX
XX 183 GTGGCTGTGACGTCAAGCAAGCCCGTGTGTTATGACATGTAAGAAACACTATCAGATAAAT 242
XX
XX 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 100
XX
XX 243 CAACAGTGGGAGCGGACCTTACCTAGGCAATGGCGTTGGTTGACTTGTATTGGAGGAAAGC 302
XX
XX 101 ArgGlyPheAsnCysGluSerIysProGluAlaGluGluThrCysPheAspIysTyrThr 120
XX
XX 303 CGAGGGTTTAACTGCGAGAGTAACTGAACTGAAAGACCTTGTGACCAAGTACACT 362
XX
XX 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProIysAspSerMetIleTrpAsp 140
XX
XX
XX 363 GGAAACACTTACCGAGTGGGTGACACTTATAGCGCTCTTAAGACTCCATGATCTGGGAC 422
XX
XX 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
XX
XX 423 TGTACCTGCTGCTGGGGCTGGCGAGGAGAGATAAGCTGTACCATCGCAAAACCGCTGCCAT 482
XX
XX 161 GluGlyGlyGlnSerTyrIysIleGlyAspThrTyrArgArgProHisGluThrGlyGly 180
XX
XX 483 GAAGGGGGTCACTCTACAAGATTGGTGCACACCTGGAGGAGACCATAGACTGGTGGT 542
XX
XX 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyIysGlyGluThrThrCysIysProIle 200
XX
XX 543 TACATGTTAGAGTGTGTGTCTTGGTAATGGAAAAGGAGATGGACCTCGAAGCCCATATA 602
XX
XX 201 AlaGluIysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
XX
XX 603 GCTGAGAGAGTGTGTTGATCATGCTGGGACTTCTATGTGTGCGAGAAACGTTGGAG 662
XX
XX 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
XX
XX 663 AAGCCCTTACCAAGCTGGATGATGTGATTGTACTTGGCTGGGAGAAAGCAGCGGACGC 722
XX
XX 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
XX
XX 723 ATCATTTCGCACTTCTAGAAATAGATGCAAGATCAGACACAAAGGACATCTATAGAAAT 782
XX
XX 261 GlyAspThrTrpSerIysIysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
XX
XX 783 GGAGACACCTGGAGCAAGAGGATATTCAGAGAAACCTGCTCCAGTGCATCTGCACAGGC 842
XX
XX 281 AsnGlyArgGlyGluTrpIysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
XX
XX 843 AACGGCCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTCTGCAGACACCATCATCGAGCGGA 902
XX
XX 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisArgProGlnPro 320
XX
XX 903 TCTGGCCCTTTCACCGATGTTGCTGCGAGCTGTTTACCACCGCAGCTTACCCCACTCT 962
XX
XX 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
XX
XX 963 CTTCCCTATGGCCACTGTGTCAAGACAGTGGTGGTGTCTACTCTGTGGGATCAGTGG 1022
XX
XX 341 LeuIysThrGlnGlyAsnIysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
XX
XX 1023 CTGAAGACACAAAGAAATAAGCAAAATGCTTTGCACGTGCTCTGGGCAACGAGTCACTGC 1082
XX
XX 361 GlnGluThrAlaValThrGlnThrTyrGlyIysAsnSerAsnGlyGluProCysValLeu 380
XX
XX 1083 CAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCCATGTCTTA 1142
XX
XX 381 ProPheThrTyrAsnAspArgThrAspSerThrThrSerAsnTyrGluGlnAspGlnIys 400
XX
XX 1143 CCATTCCACTACATGGCAGGACGTGCAGCACAACTTCGAATTATGAGCAGGACCAGAAA 1202
XX
XX 401 TyrSerPheCysThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGly 420
XX
XX 1203 TACTCTTCTGCACAGACCACTGTTTGGTTTGGTTCAGACTCGAGGAGGAGAAATTCGAATGGT 1262
XX
XX 421 AlalaCysHisPheProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGlu 440
XX
XX 1263 GCCTTGTGCCACTTCCCTTCCCTATACAACCAACCAAAATACACTGATTGCACTTCTGAG 1322
XX
XX 441 GlyArgArgAspAsnMetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnIys 460
XX
XX 1323 GGCAGAGAGACACATGAAGTGGTGGGACCAACAGAACTATGATGCCGACGAGAG 1382
XX
XX 461 PheGlyPheCysProMetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMet 480
XX
XX 1383 TTTGGTCTTCCCTCCCTGCTGCCACGAGGAAATCTGCACAACCAATGAAGGGGTCTATG 1442
XX
XX 481 TyrArgIleGlyAspGlnTrpAspIysGlnHisAspMetGlyHisMetMetArgCysThr 500
XX
XX 1443 TACCGCATTTGAGATCAGTGGGTAAGCAGCATGACATGGGTTCACATGATGAGGTGCACG 1502
XX
```

Qy 501 CysValGlyAenGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGln 520
Db 1503 TGTGTGGAAATGTCGTGGGAATGACATGTCATTGCTACTCGCAGCTCGAGATCAG 1562
Qy 521 CysIleValAspAspIleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGlu 540
Db 1563 TGCATTGTGTGATGATCATCATCTTCAATGTGAACACACATTTCCACAAAGCGTCATGAAGAG 1622
Qy 541 GlyHisMetLeuAenCysThrCysPheGlyGlnGlyArgGlyArgTyrLysCysAspPro 560
Db 1623 GGGCACATGCTGAATGTCATCATGCTTCGGTCAGGGTCGGGGCAGGTGGAGTGTGATCCC 1682
Qy 561 ValAspGlnCysGlnAspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGlu 580
Db 1683 GTCCACCAATGCCAGGATTCAGAGACTGGGACGTTTTTATCAAAATGGAGATTTCATGGAG 1742
Qy 581 LysTyrValHisGlyValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrp 600
Db 1743 AAGTATGTGCATGTTGTGATGATCAGATACCATGCTACTGCTATGCGCGTGGCATTTGGGGAGTGG 1802
Qy 601 HisCysGlnProLeuGlnThrTyrProSerSerGlyProValGluValPheIleThr 620
Db 1803 CATGCCAACCTTTACAGACCTATCCAGCTCAAGTGTGCTGTGCGAGTATTTATCACT 1862
Qy 621 GluThrProSerGlnProAsnSerHisProIleGlnTrpAenAlaProGlnProSerHis 640
Db 1863 GAGACTCCGAGTCAGCCCACTCCACCCCATCCAGTGGAAATGCACACACGCACTTCAC 1922
Qy 641 IleSerLysTyrIleLeuArgTyrArgProValSerIleProProArgAenLeuGlyTyr 660
Db 1923 ATTTCCAAGTACATCTCTCAGGTGGAGACCTGTGAGTATCCACCCAGAAACCTTTGGGATAC 1982
RESULT 5
ADQ38581
ID ADQ38581 standard; DNA; 2488 BP.
XX
AC ADQ38581;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human SNP containing myocardial infarction-associated gene, SEQ ID 244.
XX
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2004058052-A2.
XX
PD 15-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US040978.
XX
PR 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Cargill M, Devlin JJ, Iakoubova O;
XX
DR WPI; 2004-533949/51.
DR P-PSDB; ADQ39409.
XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 7; SEQ ID NO 244; 145pp; English.
XX
CC The invention relates to a novel method for identifying an individual who

CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX
SQ Sequence 2488 BP; 610 A; 641 C; 694 G; 541 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 1,76e-243 Length: 2488
Score: 3741.50 Matches: 657
Percent Similarity: 97.33% Conservative: 0
Best Local Similarity: 97.33% Mismatches: 3
Query Match: 98.96% Indels: 15
DB: 13 Gaps: 1

US-09-581-651D-1 (1-660) x ADQ38581 (1-2488)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAenMetLeu 20
Db 319 AACTTGTGTGCAACTTGGCTCCCGGTGGGGGCTCTCTCCCCACCGCTCTCAACATGCTT 378
Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
Db 379 AGGGGTCCGGGGCCCGGGCTGCTGCTGCTCGCCGCTCGCTGGGGACAGCGGTGCC 438
Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro 60
Db 439 TCCACGGGAGCCTCGAAGAGCAAGAGCAGGCTCAGCAATGGTTTCAGCCCCCAGTCCCCG 498
Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAenGlyLysHisTyrGlnIleAen 80
Db 499 GTGGCTGTCTAGTCAAGCAAGCCCGGTTGTTATGACATATGAAACACTATCATGATTAAT 558
Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysTyrGlyGlySer 100
Db 559 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTTGTTGTTGTTGTTGTTGTT 618
Qy 101 ArgGlyPheAenCysGluSerLysProGluAlaGluThrCysPheAspLysTyrThr 120
Db 619 CGAGGTTTTTAACATCGCAGAGATTAACACCTGAAGAGACTTGTGTTTGAACAATGACACT 678
Qy 121 GlyAenThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
Db 679 GGGAAACATTTACCGAGTGGGTGACATTTATGAGGCTCTTAAGACTCCATGATCTGGGAC 738
Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAenArgCysHis 160
Db 739 TGTACCTGCATYGGGGCTGGGGCGAGGAGAAATAGCTGTACCATCGCAACCGCTGCCAT 798
Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180

Db 799 GAAGGGGTAGTCCTACAAGATTGGTGACACCTGGAGGACACACATGAGACTGGTGGT 858
Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluThrPheCysLysProIle 200
Db 859 TACATGTTAGAGTGTGTGTCTTGGTAATCGAAAGAGAGAAATGGACCTCGAAGCCCATTA 918
Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTyrGlu 220
Db 919 GCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTGTGGTGGGAAACGTGGAG 978
Qy 221 LysProTyrGlnGlyTyrMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db 979 AAGCCCTACCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
Db 1039 ATCATCTTGCACTTCTAGAAAATAGATGCAACGATCAGGACACAAAGGACATCTTAGAATT 1098
Qy 261 GlyAspThrTyrPheSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 1099 GGAGACACCTGGAGCAAGAGAGATAATCGAGAAACCTGCTCCAGTGATCTGCACAGGC 1158
Qy 281 AsnGlyArgGlyGluThrLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1159 AACGGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCCATCAGCGGA 1218
Qy 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
Db 1219 TCTGGCCCTTCAACGATGTTCTGTCAGCTGTGTTACCAACCCGACGCTCACCCTCAGCCT 1278
Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTyr 340
Db 1279 CTTCCCTATGGCACTGTGTACAGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1338
Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1339 CTGAAGACACAAGGAAATTAAGCAATGCTTTGACAGTGCCTGGGCAACGGAGTCAGCTGC 1398
Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCysValLeu 380
Db 1399 CAAGAGACAGCTGTAAACCGACTTACGGTGGCACTCAAAATGGAGAGCCATGTGTCTTA 1458
Qy 381 ProPheThrTyrAsnAspArgThr----- 388
Db 1459 CCAITCACTACATGGCAGGAGCTTACTCTCTGCACACACAGAAGGGCGACAGACCGA 1518
Qy 389 -----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
Db 1519 CATCTTTGGTGCAGCACAACTTCGAATTTATGAGCAGGACCAAGAAATACTCTTTCTGCACA 1578
Qy 406 AspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyValAlaLeuCysHisPhe 425
Db 1579 GACCACACTGTTTGGTTCAGACTCGAGGAGGAAATTCCAATGGTGCCTTGTGCCACTTC 1638
Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyValArgArgAspAsn 445
Db 1639 CCCCCTCTATACACACACCAATTAACACTGATTCGACTTCTGAGGGCAGAGAGACAAAC 1698
Qy 446 MetLysTyrCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
Db 1699 ATGAAGTGGTGGGACACACAGAACTATGATGCCGACCAAGAGTTTGGGTCTGCCCC 1758
Qy 466 MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
Db 1759 ATGGCTGCCACGAGGAAATCTGCACCAACCAATGAAGGGGTTCATGTACCGCATTCGAGAT 1818
Qy 486 GlnTyrAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
Db 1819 CAGTGGGATAGCAGCAGTGCATGGGTTCATGATGAGGTGCACGTGTGTGGGAATGGT 1878
Qy 506 ArgGlyGluThrThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
Db 1879 CGTGGGAAATGACATGCAATTCCTACTCGCAGCTTCGAGATCAGTGCATTTGTTGATGAC 1938

Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545
Db 1939 ATCATTTACATGTTGAACGACACATTTCCACAGCGTCATGAAGAGGGGCACATGCTGAAC 1998
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTyrLysCysAspProValAspGlnCysGln 565
Db 1999 TGTACATGCTTCGGTCAGGGTCGGGGCAGGTGGAAGTGTGATCCCGTCGACCAATGCCAG 2058
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTyrGluLysTyrValHisGly 585
Db 2059 GATTTCAGAGACTGGGAGCTTTTATCAAAATGGAGATTTCATGGGAGAAAGTATGTGATGGT 2118
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTyrHisCysGlnProLeu 605
Db 2119 GTCAAGATACCAAGTGTCTACTGCTATGCCGTGGCATTTGGGAGTGGCATTGCCAACCTTTA 2178
Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2179 CAGACCTATCCAAGCTCAAGTGTCTGTGCAAGTATTTATCACTGAGACTCCGAGTCAG 2238
Qy 626 ProAsnSerHisProIleGlnTyrAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2239 CCCAACTCCCAACCCCATCCAGTGGATGCACCAAGCCATCTCACATTTCCAGATACATT 2298
Qy 646 LeuArgTyrArgProValSerIleProProArgAsnLeuGlyTyr 660
Db 2299 CTCAGGTGGAGACTGTGAGTATCCCAACCCAGAAACCTTGGATAC 2343

RESULT 6

ADR66637
ID ADR66637 standard; DNA; 7242 BP.
XX ADR66637;
AC ADR66637;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human prostatic carcinoma derived DNA SEQ ID 149 #3.
XX
KW human; cytostatic; diagnosis; prostatic cancer;
KW differential expression analysis; ds.
XX

OS Homo sapiens.

XX WO2004076614-A2.
XX

PD 10-SEP-2004.

XX 22-FEB-2004; 2004WO-DE000433.
XXPR 27-FEB-2003; 2003DE-01009985.
PR 14-MAY-2003; 2003DE-01022134.

XX (HINZ)/ HINZMANN B.

XX (DAHL)/ DAHL E.

XX (ROSE)/ ROSENTHAL A.

XX (HERM)/ HERMANN K.

XX (PILA)/ PILARSKY C.

XX Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarczyk C, Specht T;

PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;

PI Xinzhang L, Staub E;

XX WPI; 2004-653386/63.

DR

XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,

PT useful for diagnosis, treatment and in screening for specific binding

PT agents.

XX

PS Claim 1; Page 1040-1042; 1607pp; German.

XX This invention describes novel cytostatic polynucleotide and polypeptide

CC sequences which can be used in a method for diagnosing prostatic cancer

CC

Db 633 TGTACCTGCATCGGGCTGGCGAGGAGAAATAAGCTGTACCATCGCAAAACCGCTGCCAT 692
Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180
Db 693 GAAGGGGTGAGTCTCTACAAATGGTGACACCTGGAGGAGACACATGAGACTGGTGGT 752
Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluThrTrpCysLysProIle 200
Db 753 TACATGTTAGAGTGTGTCTTGGTAAATGGAAAAAGGAGAAATGGACCTGCAAGCCCAATA 812
Qy 201 AlaGluLysCysPheAspHisAlaIleGlyThrSerTyrValValGluGluThrTrpGlu 220
Db 813 GCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTGGTGGGAGAAACGTGGAG 872
Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluLysSerGlyArg 240
Db 873 AAGCCCTACCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 932
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
Db 933 ATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGCACATCTATAGAATT 992
Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 993 GGAGACACCTGGAGCAAGAGGATAATCGAGAAACCTGCTCCAGTGCAATCTGCACAGGC 1052
Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1053 AACGGCCGAGAGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACACACATCGAGCGGA 1112
Qy 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
Db 1113 TCTGGCCCTTCACCGATGTTCTGTCAGCTGTTTACCAACCGCAGCCTCACCCCGAGCCT 1172
Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
Db 1173 CCTCCCTATGGCCACTGTGTACAGACAGTGGTGTGGTGTCTACTCTGTGGGGATGCAGTGG 1232
Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1233 CTGAGACACAAAGGAAATAGCAAAATGCTTTGACAGTGCCTGGGCAACGGAGTCAGCTGC 1292
Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyLysAsnSerAsnGlyGluProCysValLeu 380
Db 1293 CAAGACACAGCTGTAAACCCAGACTTACGGTGGCAACTCAATGGAGAGCCATGTGTCTTA 1352
Qy 381 ProPheThrTyrAsnAspArgThr----- 388
Db 1353 CCATTCACTCAATGGCAGGACGTTCTACTCTGCACACAGAGGGCGACAGGACGGA 1412
Qy 389 ----- AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
Db 1413 CATCTTGGTCAGCAACAACCTTCGAATTATGAGCAGGACCAAGAAATCTCTTTCTGCACA 1472
Qy 406 AspHisThrValLeuValGlnThrArgGlyLysAsnSerAsnGlyAlaLeuCysHisPhe 425
Db 1473 GACCACACTGTTTGGTTTCAGACTCGAGAGGAAATTCATATGGTGGCTTGTGCCACTTC 1532
Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgArgAspAsn 445
Db 1533 CCTTCTCTATACAAACAACCAAAATACACTCAATGTCACTTCTGAGGGGAGAGAGACAAC 1592
Qy 446 MetLysTyrCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
Db 1593 ATGAAGTGGTGTGGACCAACACAGAACTATGATGCCGACCAAGATTTGGGTTCTGCCCC 1652
Qy 466 MetAlaAlaHisGluGluLeuCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
Db 1653 ATGGCTGCCACAGGAAATCTGCACAAACCAATGAAGGGGTCAATGATCCGCAATTTGGAGAT 1712
Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
Db 1713 CAGTGGGATAAGCAGCATGACATGGGTGCATGATGATGAGGTGCACGTGTGTGGGAATGGT 1772

Qy 506 ArgGlyGluThrTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
Db 1773 CGTGGGAATGGACATGCATTCCTACTCGCAGCTTCGAGATCAGTGCATTTGTTGATCAC 1832
Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysValHisGluGluGlyHisMetLeuAsn 545
Db 1833 ATCACTTACAAATGTGAACGACACATTCACAAAGCGTCATGAAGAGGGGCACATGCTGAAC 1892
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArg 565
Db 1893 TGTACATGCTTCGGTCAAGGTCGGGGCAGGTGGAAAGTGTGATCCCGTCGACCAATGCCAG 1952
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTyrGluLysTyrValHisGly 585
Db 1953 GATTACAGAGACTGGAGCGTTCATCAAAATGGAGATTCATGGGAGAGTATGTGTCATGGT 2012
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluThrHisCysGlnProLeu 605
Db 2013 GTCAGATACCACTGCTACTGCTATGGCGTGGCAATGGGGAGTGGCAATGCCAACCTTTA 2072
Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2073 CAGACCTATCAAGCTCAAGTGGTCTCTGTCGAAGTATTTATCACTGAGACTCCGAGTCCG 2132
Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2133 CCCAACTCCCAACCCCTCCAGTGGATGCACACAGCCATCTCACATTTCCCAAGTACATT 2192
Qy 646 LeuArgTrpArgPro 650
Db 2193 CTCAGGTGGAGACCT 2207
RESULT 8
ACC00412
ID ACC00412 standard; cDNA; 7550 BP.
XX
AC ACC00412;
XX
DT 04-JUL-2003 (first entry)
XX
DE Human cell adhesion and extracellular matrix protein, CADECM-21, DNA.
XX
KW Human; anti-HIV; anti-allergic; cerebroprotective; antiparkinsonian;
KW anticonvulsant; nootropic; neuroprotective; immunosuppressive;
KW dermatological; anti-inflammatory; cytostatic; antiarteriosclerotic;
KW gene therapy; cell adhesion; extracellular matrix; CADECM;
KW immune system disorder; AIDS; allergy; neurological disorder; stroke;
KW parkinson's disease; epilepsy; developmental disorder; Down's syndrome;
KW cerebral palsy; connective tissue disorder; systemic lupus erythematosus;
KW genetic disorder; Alport's syndrome; cell proliferative disorder; cancer;
KW atherosclerosis; gene; 88.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 271..7071
FT /*tag= a
FT /product= "CADECM-21"
XX
XX WO2003027230-A2.
XX
XX 03-APR-2003.
XX
XX 02-AUG-2002; 2002WO-US024649.
XX
XX 03-AUG-2001; 2001US-0309964P.
XX 03-AUG-2001; 2001US-0310119P.
XX 17-AUG-2001; 2001US-0313091P.
XX 31-AUG-2001; 2001US-0316771P.
XX 07-SEP-2001; 2001US-0317896P.
XX 21-SEP-2001; 2001US-0324781P.
XX 05-OCT-2001; 2001US-0327606P.

PR 12-OCT-2001; 2001US-0328960P.
PR 09-NOV-2001; 2001US-0344471P.
PR 17-MAY-2002; 2002US-0381291P.
XX (INCY-) INCYTE GENOMICS INC.
XX Burford N, Warren BA, Duggan BM, Mason PM, Richardson TW, Yue H;
PI Forsythe IJ, Elliott VS, Griffin JA, Gorvad AE, Azimzai Y;
PI Kallick DA, Xu Y, Honchell CD, Baughn MR, Gietzen KJ, Lee S;
PI Walia NK, Tang YT, Nguyen DB, Becha SD, Lee SY, Ramkumar J;
XX WPI; 2003-354645/33.
DR P-ESDB; ABR40124.
XX
XX New human cell adhesion and extracellular matrix proteins (CADECM),
PT useful for diagnosing, treating or preventing disorders associated with
PT aberrant CADECM expression, e.g. cancer, AIDS, atherosclerosis, allergies
PT or stroke.
XX
XX Claim 12; Page 231-233; 234pp; English.
XX
XX The present invention relates to novel human cell adhesion and
CC extracellular matrix proteins, CADECM-1 to CADECM-22, and their coding
CC sequences (ACC00392-ACC00413 and ABR40104-ABR40125). The coding sequences
CC and proteins are useful in diagnosing, treating and preventing disorders
CC associated with aberrant expression of CADECM, such as immune system
CC disorders (e.g. AIDS or allergies), neurological disorders (e.g. stroke,
CC Parkinson's disease or epilepsy), developmental disorders (e.g. Down's
CC syndrome or cerebral palsy), connective tissue disorders (e.g. systemic
CC lupus erythematosus), genetic disorders (e.g. Alport's syndrome) or cell
CC proliferative disorders (e.g. cancer or atherosclerosis)
XX
SQ Sequence 7550 BP; 1996 A; 1993 C; 1889 G; 1672 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.15e-239 Length: 7550
Score: 3693.50 Matches: 648
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 2
Query Match: 97.69% Indels: 15
DB: 8 Gaps: 1

US-09-581-651D-1 (1-660) x ACC00412 (1-7550)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 217 AACTTGGTGGACACTTGGCTCCCGTGGGGGTCTCTCCCCACCGCTCAACATGCTT 276

Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
Db 277 AGGGTCCGGGGCCGGGCTGCTGCTGGCCGTCAGTGCCTGGGACAGCGGTGCC 336

Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro 60
Db 337 TCCACGGAGGCTCGAAGAGCAAGAGGAGGCTCAGCAATGGTTCAGCCCCAGTCCCG 396

Qy 61 ValAlaValSerGlnSerLysProGlyCysThrAspAsnGlyLysHisThrGlnIleAsn 80
Db 397 GTGGCTGTGCTCAAGACAGCCCGGTGTATGACATGGAAACACTATCAGATAAT 456

Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 100
Db 457 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGGAGGAAGC 516

Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluThrCysPheAspLysThrThr 120
Db 517 CGAGGTTTAACTGCGCAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTGCACAACTAC 576

Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
Db 577 GGAACACTTACCGAGTGGGTGACACTATATGAGCGCTCTTAAGACTCCATGATCTGGAC 636

Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160

Db 637 TGTACCTGTCATCGGGGCTGGCGGAGGAGAATAAGCTGTACCATCGCAAAACCGCTGCAT 696
Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180
Db 697 GAAGGGGGTCACTCTCAAGATTGGTGACACCTCGGAGGAGACCAATAGACTGTGGT 756
Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysPhePro 200
Db 757 TACATGTTAGAGTGTGTGCTCTTGGTAATGGAAAGAGAGATGGACCTCGCAAGCCCAT 816
Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
Db 817 GCTGAGAAGTGTTTGTATCATCTGCTGGGACTTCTTATGTGGTGGAGAAACCTGGAG 876
Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db 877 AGCCCTTACCAGGCTGGATGATGTAGATGTACTTGCCTGGGAAAGGAGGAGGAGCGC 936
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
Db 937 ATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTATAGAAT 996
Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 997 GGAGACACCTGGAGCAAGAGGATTAATCGAGGAACCTGCTCCAGTGCATCTGCACAGGC 1056
Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1057 AACGGCCGAGAGAGTGGAGTGTGAGAGGACACCTCTGTGCAGACACACATCGAGCGGA 1116
Qy 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
Db 1117 TCTGGCCCTTCCCGATGTTGTCAGCTGTTTACCACCGCAGCCTCACCCCCAGCCT 1176

Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValValTyrSerValGlyMetGlnTrp 340
Db 1177 CCTCCTATGCGCACTGTGTGCACAGACAGTGGTGTGCTACTCTGTGGGGATCGAGTGG 1236

Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1237 CTGAAGACACAGGAAATAGCAANTGCTTTGCAGCTGCTGGGCAAGGAGTCACTGC 1296

Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyLysAsnSerAsnGlyGluProCysValLeu 380
Db 1297 CAAGAGACAGCTGTAACCCAGACTTACGCTGGCAACTCAATGGAGAGCCATGTGTCTTA 1356

Qy 381 ProPheThrTyrAsnAspArgThr----- 388
Db 1357 CCATTCACTACAAATGGCAGGACGCTTCTACTCTCCACCAAGAGGCGGACAGACGGA 1416

Qy 389 -----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
Db 1417 CATCTTTGGTGCACACAACTTCGAANTATGAGCAGGACCAAGAAATACTCTTTCTGCACA 1476

Qy 406 AspHisThrValLeuValGlnThrArgGlyLysAsnSerAsnGlyAlaLeuCysHisPhe 425
Db 1477 GACCACACTGTTTGGTTTCAGACTCGAGGAGAAATCCAAATGCTGCTGTCCTTGTGCCACTTC 1536

Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgAspAsn 445
Db 1537 CCTTCTCTATACAAACCACTTACACTGATTTGCACTTCTCGAGGGCAGAGAGACAAAC 1596

Qy 446 MetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
Db 1597 ATGAAGTGGTGGGACCAACACAGAACTATGATGCCGACCAAGATTTGGGTTCTGCCCC 1656

Qy 466 MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
Db 1657 ATGGCTGCCACGAGGAAATCTGCACAACTCAATGAAGGGGTCTATGTATCCGCTTGGAGAT 1716

Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505

Db 1717 CAGTGGGATAGGACGATGATGGTTCACATGATGAGGTGCAGCTGTGGGAATGGT 1776
Qy 506 ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
Db 1777 CGTGGGAATGGACATGATGCTTCCCTACTCGCAGCTTCGAGATCAGTCATTTGTATGAC 1836
Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545
Db 1837 ATCACTTACAAATGGAACGACACATTCACAAGCGCTCATGAAGAGGGGACATGCTGAAC 1896
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 565
Db 1897 TGTACATGCTTCGTCAGGCTCGGGGAGGTGGAGTGTATCCCTCGACCAATGCCAG 1956
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluTyrValHisGly 585
Db 1957 GATTACAGAGCTGGGACGTTTATCAAAATGGAGATTCATGGGAGAAATGATGTGATGGT 2016
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu 605
Db 2017 GTCAGATACCAGTGTACTGCTATGGCGGTGGCATTTGGGAGTGGCATTCGCAACCTTTA 2076
Qy 606 GlnThrTyrProSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2077 CAGACCTATCCAGCTCAAGTGTCTCTGCGAGTATTTATCACTGAGACTCCGAGTCAG 2136
Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2137 CCCAACTCCCAACCCATCCAGTGAATGCACACAGCCATCTCACATTTCCNAGTACATT 2196
Qy 646 LeuArgTrpArgPro 650
Db 2197 CTCAGGTGGAGACCT 2211
RESULT 9
ADP64998
ID ADP64998 standard; DNA; 8027 BP.
XX
AC ADP64998;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human fibronectin 1 (FNI), transcript variant 1 DNA sequence.
XX
KW autoimmune disease; arthritis; gene expression analysis;
KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; ds; human.
XX
OS Homo sapiens.
XX
PN W02003072827-A1.
XX
XX
PD 04-SEP-2003.
XX
XX 31-OCT-2002; 2002WO-US035433.
XX
XX 31-OCT-2001; 2001US-0336220P.
XX
XX (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
PA
XX
PI Hirsch R, Thorton SL;
XX
XX WPI; 2003-712740/67.
DR GENBANK; NM_002026.
XX
XX
PT Diagnosing and analyzing autoimmune disease using gene expression
PT profiles and microarray technology, useful for diagnosing and treating
PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT gout.
XX

PS Disclosure; Page; 56pp; English.

XX The invention relates to a novel method for diagnosing and analysing
CC autoimmune disease or arthritides. The method comprises obtaining a
CC patient sample containing mRNA, analysing gene expression using the mRNA
CC that results in a gene expression signature of the mRNA, and using that
CC gene expression signature to diagnose or analyse the autoimmune disease
CC or arthritides in the patient, where gene expression of at least 60% of
CC the genes correlates with that of the gene signature. The invention
CC further comprises: a treatment of rheumatoid arthritis; identification of
CC genes for targeting in the treatment of rheumatoid arthritis in a mammal
CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
CC array or gene chip, specific for rheumatoid arthritis; diagnosis or
CC analyses of autoimmune disease or rheumatoid arthritis; screening the
CC efficacy of a candidate drug in vitro for the treatment of collagen-
CC induced arthritis; and reducing the symptoms associated with collagen-
CC induced arthritis. The compositions of the invention have the following
CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
CC antigout, antiinflammatory, dermatological, and immunomodulatory. The
CC methods and compositions of the present invention are useful for
CC diagnosing and treating autoimmune disease or arthritides, such as
CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
CC immune disease caused by an infectious agent. This polynucleotide
CC represents a DNA sequence relating to the genes used in the analysis and
CC treatment of autoimmune diseases or arthritides. Note: This sequence is
CC not shown in the specification. It has been supplied in an electronic
CC format from WIPO.
XX
SQ Sequence 8027 BP; 2130 A; 2111 C; 1973 G; 1813 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.24e-239 Length: 8027
Score: 3693.50 Matches: 648
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 2
Query Match: 97.69% Indels: 15
DB: 11 Gaps: 1

US-09-581-651D-1 (1-660) x ADP64998 (1-8027)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 214 AACTTGGTGGCAACTTGCCTCCCGGTGCGGGCGCTCTCTCCCCACCGCTCTCAACATGCTT 273
Qy 21 ArgGlyProGlyProGlyLeuLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
Db 274 AGGGGTCCGGGGGCCGGGGTGTCTGTCTGGCCGCTCCAGTGCCTGGGGACAGCGGTGCC 333
Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
Db 334 TCCACGGGAGCCTCGAAGAGCAAGAGCAGGCTCAGCAATGGTTTCAGCCCGCCAGTCCCG 393
Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnIleAsn 80
Db 394 GTGGCTGTCTAGTCAAGCAAGCCCGGTGTGTATGACAATGGAAACACATATCATGATAAAT 453
Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 100
Db 454 CAACAGTGGGAGCGGACCTACCTAGGCATGCGTGTGGTTTGTACTTGTATGGAGGAGC 513
Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysTyrThr 120
Db 514 CGAGGTTTTTAACCTCGAGAGTAGTAACCTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACT 573
Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
Db 574 GGGAACACTTACCAGGTGGGTGACACTTATGACGCTCTTAAGACATCCATGATCTGGGAC 633
Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
Db 634 TGTAACCTGCAATCGGGCTGGCGAGGGAGAGATTAAGCTGTACCATCGCAACCCGCTGCCAT 693

Qy	161	GlulGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly	180
Db	694	GNAGGGGGTCAGTCTCTCAAGATTGGTGACACCTGGAGGAGACACATGAGACTGGTGGT	753
Qy	181	TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluThrTrpThrCysLysProIle	200
Db	754	TACATGTTAGTAGTGTGTGTTGGTAATGGAAAGAGGAATGGACCTCGAAGCCCAT	813
Qy	201	AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu	220
Db	814	GCTGAGAAGAGTTTTGATCATGCTCTGGGACTTCCTATGTGTGCGAGAAACGTGGGAG	873
Qy	221	LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg	240
Db	874	AAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCTGGGAGAAGGCAGCGGACGC	933
Qy	241	IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle	260
Db	934	ATCACTTTGCACCTCTAGAAATAGATGCACACGATCAGGACACAAGGACATCCTATAGAATT	993
Qy	261	GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysThrGly	280
Db	994	GGAGACACTGGAGCAAGAAGATATTCGAGAAACCTGCTCCAGTGCATCTGCGACAGGC	1053
Qy	281	AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly	300
Db	1054	AACGGCCGAGGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCGAGCGGA	1113
Qy	301	SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro	320
Db	1114	TCTGGCCCTTTCACCGATGTTTCGTGAGCTGTTTACCAACCGCAGCCTCACCCCCAGCCT	1173
Qy	321	ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp	340
Db	1174	CCTCCCTATGGCCACTGTGTCAAGACAGTGGTGTGTCTACTCTGTGGGGATGCGATGG	1233
Qy	341	LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys	360
Db	1234	CTGAAGACACAANGAAATAAGCAAAATGCTTTGCACGTGCTGGGCAACGGAGTCAGCTGC	1293
Qy	361	GlnGluThrAlaValThrGlnThrTyrGlyLysAsnSerAsnGlyGluProCysValLeu	380
Db	1294	CAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTCTCTTA	1353
Qy	381	ProPheThrTyrAsnAspArgThr-----	388
Db	1354	CCATTCACCTACAATGGCAGAGACGTCTACTCTCTGCACCAACGGAGGGCGACAGACGGA	1413
Qy	389	-----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr	405
Db	1414	CATCTTTGGTCGACACACAACCTTCGAATTTATGACGAGGACACAGAAATACTCTTTCTGCACA	1473
Qy	406	AspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCysHisPhe	425
Db	1474	GACCACACTGTTTGGTTTCAGACTCGAGAGAGAAATTCCAATGGTGGCTTGTGGCACCTTC	1533
Qy	426	ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerSerGluGlyArgArgAspAsn	445
Db	1534	CCCTTCCTATACAACAACCAATTTACACTGATTTGCATCTTCTGAGGGCAGAAGAGACAAAC	1593
Qy	446	MetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro	465
Db	1594	ATGAAGTGGTGTGGGACACACAGAACTATGATCGGACACAGAGTTTGGGTTCGCCCC	1653
Qy	466	MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp	485
Db	1654	ATGGTGGCCACGAGGAATCTGCAACAACCAATTAAGGGGGTCATGTACCGCATTTGGAGAT	1713
Qy	486	GlnTrpAspLysGlnHisAspMetClyHisMetMetArgCysThrCysValGlyAsnGly	505
Db	1714	CAGTGGGATAGCAGCATGACATGGGTCAACATGATGAGGTGCACGTGTGTGGGAATGGT	1773
Qy	506	ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp	525

Db	1774	CGTGGGAAATGGACATGCATTCGCTACTCCAGCTTCGAGATCAGTGCAATGTTGATGAC	1893
Qy	526	IleThrTyraenValaenAepThrPheHisLysAraHisGluGluGlyHisMetLeuAen	545
Db	1834	ATCATTACAAATGTGAACACACATCCACNAGCGTCATGAAGAGGGGCACATGCTGAAC	1893
Qy	546	CyethrCysPheGlyGlnGlyArgGlyArgTrrPlyCysAepProValAepGlnCysGln	565
Db	1894	TGTACATGCTTCGGTCAGGGTCGGGGCAGGTGGAAGTGTATCCCGTCGACCAATGCCAG	1953
Qy	566	AspSerGluThrGlyThrPheTyrrGlnIleGlyAepSerTrpGluLysTyrrValHisGly	585
Db	1954	GAITTCAGAGACTCGGACGTTTATCAAAATGGAGATTCACTGGGAGAAAGTATGTCATGGT	2013
Qy	586	ValArgTyrrGlnCysTyrrCysTyrrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu	605
Db	2014	GTCAATACACAGTGTACTCTCTATGGCCGTGGCATTGGGGAGTGGCATTGCCAACCTTTA	2073
Qy	606	GlnThrTyrrProSerSerGlyProValGluValPheIleThrGluThrProSerGln	625
Db	2074	CAGACCTATCCAAAGCTCAAGTGTCTCTGTCGAAGTATTTATCACTGAGACTCCGAGTCAG	2133
Qy	626	ProAenSerHisProIleGlnTrpAenAlaProGlnProSerHisIleSerLysTyrrIle	645
Db	2134	CCCAACTCCCAACCCCATCCAGTGGAAATGCACCACAGCCATCTCACATTTCCAAGTACATT	2193
Qy	646	LeuArgTrpArgPro	650
Db	2194	CTCAGGTGGAGACCT	2208
RESULT 10			
ADG89565			
ID	ADG89565 standard; DNA; 8027 BP.		
XX	ADG89565;		
AC			
XX			
DT	11-MAR-2004 (first entry)		
XX	Human fibronectin gene #3.		
DE			
XX	metastatic cancer cell differentiation; mutated fibronectin;		
KW	metastatic cancer; human; ds; gene; fibronectin.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	268..7335	
FT		/*tag= a	
FT		/product= "Human fibronectin protein"	
XX	WO2003100027-A2.		
PN			
XX			
PD	04-DEC-2003.		
PF	28-MAY-2003; 2003WO-US016736.		
XX			
PR	28-MAY-2002; 2002US-0383530P.		
XX			
FA	(BAYU) BAYLOR COLLEGE MEDICINE.		
XX			
PI	Wang R;		
XX			
DR	WPI; 2004-035134/03.		
DR	P-FSDB; ADG89560.		
XX			
PT	Identifying a cell that differentiates into a metastatic cancer cell,		
PT	useful for preventing metastatic cancer, comprises identifying a mutated		
PT	fibronectin in the cell.		
XX			
PS	Disclosure; SEQ ID NO 8; 137pp; English.		
XX	The invention comprises a method for identifying a cell that will		

CC differentiates into a metastatic cancer cell, the method involves
CC identifying a mutated fibronectin in the cell. The method of the
CC invention is useful for preventing metastatic cancer. The present DNA
CC sequence represents a human fibronectin gene.

XX
SQ Sequence 8027 BP; 2130 A; 2111 C; 1973 G; 1813 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.24e-239 Length: 8027
Score: 3693.50 Matches: 648
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 2
Query Match: 97.63% Indels: 15
DB: 12 Gaps: 1

US-09-581-651D-1 (1-660) x ADG89565 (1-8027)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 214 AACTTGGTGGCAACTTGCCTCCCGTGGGGCGTCTCTCCCCACCGCTCTCAACATGCTT 273
Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
Db 274 AGGGTCCGGGGCCCGGGCTGCTGCTGCTGGCCGTCAGTGCCCTGGGGACAGCGGTGCC 333
Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
Db 334 TCCACGGGAGCCTCGAAGAGCAGAGGAGGAGCTCAGCAAAATGGTTCCGCCCCAGTCCCG 393
Qy 61 ValAlaValSerGlnSerLysProGlyCysThrAspAsnGlyLysHisThrGlnIleAsn 80
Db 394 GTGGCTGTGCTCAAGCAAGCCGGTGTGTATGACAAATGGAAACACATATCAGATAAT 453
Qy 81 GlnGlnTrpGluArgThrLysLeuGlyAsnAlaLeuValCysThrCysThrGlyLysSer 100
Db 454 CAACAGTGGGAGCGGACCTACTAGGCAATCGCTGGTTGTTGTTGTTATGGAGGAGC 513
Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGlnThrCysPheAspLysThr 120
Db 514 CGAGGTTTAACTCGAGAGTAAACCTGAAGCTGAAGAGACTGCTTCTTGAACAAGTACACT 573
Qy 121 GlyAsnThrThrArgValGlyAspThrThrGluArgProLysAspSerMetIleThrAsp 140
Db 574 GGGAAACCTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGAC 633
Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
Db 634 TGTACCTGCATCGGGGCTGGGCGAGGAGGAGATAAGCTGTACCATCGCAAAACCGCTGCCAT 693
Qy 161 GluGlyGlyGlnSerThrLysIleGlyAspThrThrArgArgProHisGluThrGlyGly 180
Db 694 GAAGGGGTGAGTCTCAACAGATTGGTGACACCTGGAGGAGACACCATGAGACTGGTGT 753
Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluThrThrCysLysProIle 200
Db 754 TACATGTTAGAGTGTGTGTCTTGGTAATGGAAAGGAGAAATGGAGCTGCAAGCCCAT 813
Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerThrValValGlyGluThrTrpGlu 220
Db 814 GCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCTTATGTGGTGGGAAACGTGGAG 873
Qy 221 LysProThrGlnGlyThrMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db 874 AAGCCCTACCAAGCTGGATGATGATGATGTTGACTTGCCTGGGAGAGGAGCGGAGCG 933
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerThrArgIle 260
Db 934 ATCACTTGCACTTCTAGAAAATAGATGCAACCGATCAGGACACACAGGACATCTATAGAAT 993
Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 994 GGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGCACAGCG 1053

Qy 281 AsnGlyArgGlyGluThrLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1054 AACGGCCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACACACATCGAGCGGA 1113
Qy 301 SerGlyProPheThrAspValArgAlaAlaValThrGlnProGlnProHisProGlnPro 320
Db 1114 TCTGGCCCTTACCAGATGTTGTCGAGCTGTTTACCACCGCAGCTTACCCCGAGCT 1173
Qy 321 ProProThrGlyHisCysValThrAspSerGlyValValThrValThrSerValGlyMetGlnTrp 340
Db 1174 CTTCCCTATGGCCACTGTGTGCAGACAGTGGTGGTGTCTACTCTGTGGGGATGCAGTGG 1233
Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1234 CTGAAGACACAAAGGAAATAAGCAAAATGCTTTGGACGTCCTGGGCAACGGAGTCAGCTGC 1293
Qy 361 GlnGlnThrAlaValThrGlnThrThrGlyGlyAsnSerAsnGlyGluProCysValLeu 380
Db 1294 CAAGAGACAGCTGTAACCCAGACTTACGGTGGCACTCAATATGGAGAGCCATGTGTCTTA 1353
Qy 381 ProPheThrThrAsnAspArgThr----- 388
Db 1354 CCATTCACTACAATGCCAGGACGTTCTACTCTGTCACACCGAAGGCGGACAGCGGA 1413
Qy 389 -----AspSerThrThrSerAsnThrGluGlnAspGlnLysThrSerPheCysThr 405
Db 1414 CATCTTTGGTGCAGCACAACTTCGAATATATGAGCAGGACAGCAAAATACTCTTTCTGCACA 1473
Qy 406 AspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCysHisPhe 425
Db 1474 GACCACACTGTTTGGTTCAGACTCGAGGAGGAAATTCGAATGTGTGCTTGTGCCACTTC 1533
Qy 426 ProPheLeuThrAsnAsnHisAsnThrThrAspCysThrSerGluGlyArgArgAspAsn 445
Db 1534 CCCTTCTATACAACACCACTTACTGATTTGCTCTTCTGAGGCGAGAGAGACAC 1593
Qy 446 MetLysTrpCysGlyThrThrGlnAsnThrAspAlaAspGlnLysPheGlyPheCysPro 465
Db 1594 ATGAAGTGGTGTGGGACACACAGAACTATGATGCCGACCAAGATTTGGGTTCTGCCCC 1653
Qy 466 MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetThrArgIleGlyAsp 485
Db 1654 ATGGCTCCCAAGGAGAAATCTGCACAACTATGAAGGGGTGATGATCCCGCATTTGGAGAT 1713
Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
Db 1714 CAGTGGATGAAGCAGCATGATGGGTCACTGATGAGTGCAGCTGTGTGGGAATGTT 1773
Qy 506 ArgGlyGluThrThrCysIleAlaThrSerGlnLeuArgAspGlnCysIleValAspAsp 525
Db 1774 CGTGGGAAATGGACATGCAATTCCTACTCGCAGCTTCGAGATCAGTGTGTTGTATGATGAC 1833
Qy 526 IleThrThrAsnValAsnAspThrPheHisArgHisGluGluGlyHisMetLeuAsn 545
Db 1834 ATCACTTACAAATGGAACGACACATCCCAAGCGTCATGAAGAGGGGACATCTCTGAAC 1893
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 565
Db 1894 TGTACATGCTTCGGTCAAGGTCGGGCGAGGTGGAGTGTGATCCCGTCGACCAATGCCAG 1953
Qy 566 AspSerGluThrGlyThrPheThrGlnIleGlyAspSerTrpGluLysThrValHisGly 585
Db 1954 GATTCAGAGACTGGACGTTTTTATCAAAATGGAGATTCATGGGAGAGATGATGTGATGTT 2013
Qy 586 ValArgThrGlnCysThrCysThrGlyArgGlyIleGlyGluThrPheCysGlyProLeu 605
Db 2014 GTCAATACCAAGTGTCTGCTATGGCCGTGGCATTCGGGAGTGGCATTCGCAACCTTTA 2073
Qy 606 GlnThrThrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2074 CAGACCTATCAAGCTCAAGTGTCTGTCGAGATTTTATCACTGAGACTCCGAGTCAG 2133
Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysThrIle 645

```
Db 2134 CCCAACTCCCAACCCCACTCCAGTGGAGTGCACCAAGCCATCTCAATTTCCAGTACATT 2193
Qy 646 LeuArgTrpArgPro 650
Db 2194 CTCAGGTGGAGACCT 2208

RESULT 11
ADQ29601
ID ADQ29601 standard; DNA; 8027 BP.
AC AC
XX AC
XX ADQ29601;
DT 07-OCT-2004 (first entry)
XX
XX Human colorectal cancer-associated protein coding sequence #23.
DE
XX
XX human; colon cancer; TIMP1; RegI-alpha;
XX colorectal cancer-associated marker; gene; ds.
XX
OS Homo sapiens.
XX
XX EP1439393-A2.
XX
XX 21-JUL-2004.
XX
XX 15-DEC-2003; 2003EP-00257868.
XX
XX 13-DEC-2002; 2002US-0433554P.
XX
XX 31-JUL-2003; 2003US-0491397P.
XX
XX (FARB ) BAYER HEALTHCARE LLC.
XX
XX (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.
XX
XX Astle JH, Boardman LA, Bugart LJ, Burgess CC, Catino TJ;
XX Dwivedi P, Huntress M, Johnson KA, Lewis ME, Maimonis PJ, Myerow SH;
XX Brown-Shimer SLA, Thiagalingam A, Thibodeau SN, Molino GA;
XX
XX WPI; 2004-545561/53.
XX
XX P-PSDB; ADQ29668.
XX
XX
XX Diagnosing colon cancer in individual, preferably human, by detecting
XX presence of TIMP 1 in sample, where presence of TIMP 1 in sample is
XX indicative of colon cancer in individual.
XX
XX Claim 7; SEQ ID NO 27; 433pp; English.
XX
XX The invention comprises a method for diagnosing colon cancer in an
XX individual, the method involves obtaining a serum sample from the
XX individual and detecting the presence of either TIMP1 or RegI-alpha and
XX an additional colorectal cancer-associated marker. The method of the
XX invention is useful for diagnosing colon cancer in an individual. The
XX present DNA sequence represents a human colorectal cancer-associated
XX protein coding sequence of the invention.
XX
XX
XX Sequence 8027 BP; 2130 A; 2111 C; 1973 G; 1813 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.24e-239 Length: 8027
Score: 3693.50 Matches: 648
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 2
Query Match: 97.69% Indels: 15
DB: 12 Gaps: 1

US-09-581-651D-1 (1-660) x ADQ29601 (1-8027)

Qy 1 AnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 214 AACTGGTGGCAACTGGCTCCGGTGGCGGCTCTCCCCCAACCGTCTCAACATGCTT 273
Qy 21 ArgGlyProGlyLeuLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
```

```
Db 274 AGGGTCCGGGGCCGGCTGCTGCTGCGCGCTCCAGTGCCTGGGGACAGCGGTGCC 333
Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
Db 334 TCACCGGGAGCCTCGAAGAGCAAGAGCGAGCGCTCAGCAAAATGGTTACGCCCACTCCCG 393
Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnIleAsn 80
Db 394 GTGGCTGTCACTCAAGCAAGCCCGGTTGTTATGACAATCGAAAACACATCATAGATAAT 453
Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 100
Db 454 CAACAGTGGGAGCGGACCTACCTAGSCAATGCGTTGGTTTGTACTTGTATTAGGAGAAC 513
Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysTyrThr 120
Db 514 CGAGGTTTAACTGCGAGAGTAACCTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACT 573
Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
Db 574 GGGNACACTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGAC 633
Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
Db 634 TGTACCTGCATCGGGGCTGGCGAGGAGAAATAAGCTGTACCATCGCAAAACCGCTGCAT 693
Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180
Db 694 GAAGGGGGTCACTCCTCAAGATTGGTGACACCTGGAGGAGACCACATGAGACTGTGGT 753
Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysProIle 200
Db 754 TACATGTTAGAGTGTGTGTCTTGGTAATGGAAAAGAGATGGACCTCGCAAGCCCAT 813
Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
Db 814 GCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCTCTATGTGGTGGAGAAACGTGGAG 873
Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db 874 AAGCCCTACCAAGGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 933
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
Db 934 ATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACACAGGACATCTATAGAATT 993
Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 994 GGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC 1053
Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1054 AACGGCGAGAGAGTGGAAAGTGTGAGAGGCGACACCTCTGTGCAAGACCACATCGAGCGGA 1113
Qy 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
Db 1114 TCTGGCCCTTCACCGATGTTCTGTGCGAGCTGTTTACCAACCGCGACCTCACCCCGAGCT 1173
Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
Db 1174 CCTCCCTATGGCCACTGTGTACAGACAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1233
Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1234 CTGAAGACACAAGGAAATAAGCAAAATGCTTTGACGTCGCTGGGCAACCGAGTCTAGCTGC 1293
Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyAsnSerAsnGlyGluProCysValLeu 380
Db 1294 CAAGAGACAGCTGTAAACCCAGACTTACGGTGGCACTCAATGGAGAGCCATGTGTCTTA 1353
Qy 381 ProPheThrTyrAsnAspArgThr----- 388
Db 1354 CCATTACCTACAATGGCAGGAGCTTCTACTCTCTGCAACCGAAGGCGGACAGGACGGA 1413
```

Qy 389 -----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
 Db 1414 CATCTTTGGTGCAGCACCACTTCGAATATATGAGCAGGACCAAGAAATCTCTTTCTGCACA 1473
 Qy 406 AspHisThrValLeuValGlnThrArgGlyCysAsnSerAsnGlyAlaLeuCysHisPhe 425
 Db 1474 GACCACACTGTTTGGTTCAGACTCGAGGAGGAATTCGAATGTTGCTTGTGGCCACTTC 1533
 Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgArgAspAsn 445
 Db 1534 CCCTTCCTATACACACCACTTACACTGATTCGACTTCTGAGGGCAGAGAGACAC 1593
 Qy 446 MetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
 Db 1594 ATGAAGTGGTGTGGGACCAACAGAACTATGATGCCAGCACGAAGATTGGGTTCTGCCCC 1653
 Qy 466 MetAlaAlaHisGluGluLeuLeuThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
 Db 1654 ATGGCTGCCACGAGGAATCTGCACCAACCAATGAAGGGGTCATGTACCGCATTTGGAGAT 1713
 Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
 Db 1714 CAGTGGATAGCAGCATGACATGGTCACATGATGAGTGCAGTGTGTGGGAATGGT 1773
 Qy 506 ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
 Db 1774 CGTGGGAATGCACATGCTGCTACTCGCAGCTTCGAGATCAGTGCATTTGTTGATGAC 1833
 Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545
 Db 1834 ATCACTTACATATGTAACGACACATTCACAAAGCGTCATGAAGAGGGGCACATGCTGAAC 1893
 Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 565
 Db 1894 TGTACATGCTTCGGTTCAGGGTCGGGGCAGGTGGGAAGTGTATCCCGTCGACCAATGCCAG 1953
 Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrValHisGly 585
 Db 1954 GATTACAGAGACTGGGCGCTTTTATCAAAATTCGAGATTCATGGGAGAAATATGTCATGGT 2013
 Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu 605
 Db 2014 GTCAATACCACTGCTACTGCTATGGCGGTGGCATTTGGGAGTGGCATTTGCCAACCTTTA 2073
 Qy 606 GlnThrTyrProSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
 Db 2074 CAGACCTATCCAGCTCAAGTGGTCTGTGCAAGTATTTATCACTGAGACTCCGAGTCAG 2133
 Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
 Db 2134 CCCAACTCCCAACCCATCCAGTGGAAATGCACACAGCCATCTCACATTTCCAAAGTACATT 2193
 Qy 646 LeuArgTrpArgPro 650
 Db 2194 CTCAGGTGGAGACCT 2208
 RESULT 12
 ADR67200
 ID ADR67200 standard; DNA; 8027 BP.
 XX
 AC ADR67200;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Human bladder cancer associated nucleotide sequence.
 XX
 KW bladder cancer tissue; bladder cancer; cytostatic; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN W02004076613-A2.
 XX

PD 10-SEP-2004.
 XX
 PF 24-FEB-2004; 2004WO-DE000364.
 XX
 PR 26-FEB-2003; 2003DB-01009729.
 XX
 PA (HERR/) HERR A.
 PA (HINZ/) HINZMANN B.
 PA (DAHL/) DAHL E.
 PA (STAU/) STAUB E.
 PA (PILA/) PILARSKY C.
 PA (SPEC/) SPECHT T.
 XX
 PI Herr A, Hinzmann B, Dahl E, Staub E, Pilarsky C, Specht T;
 XX WPI; 2004-653385/63.
 XX
 PT New nucleic acids, and encoded proteins, from bladder cancer tissue,
 PT useful for diagnosis, treatment and in screening for specific binding
 PT agents.
 XX
 PS Claim 1; Fig 3; 112pp; German.
 XX
 CC The present invention describes nucleic acids (I) associated with bladder
 CC cancer tissue. Also described: (1) peptides and proteins (II) containing
 CC an amino acid sequence encoded by (I); (2) a method for diagnosing
 CC bladder cancer (BC), or monitoring its progression, that uses (I), (II)
 CC or agents (2) that inhibit, or bind to, (I) or (II), labelled with a
 CC reporter; and (3) a method for treating BC that uses (I), (II) or (2).
 CC (I) and (II) have cytostatic activity. (I) and (II) can be used to detect
 CC (and monitor progression of) bladder cancer (BC), or the risk of
 CC developing it; to screen for specific binding agents (2), and to treat
 CC BC. (2) are also useful as diagnostic and therapeutic agents. The present
 CC sequence represents a human nucleotide sequence associated with bladder
 CC cancer, which is used in the exemplification of the present invention.
 XX
 SQ Sequence 8027 BP; 2130 A; 2111 C; 1973 G; 1813 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1-24e-239 Length: 8027
 Score: 3693.50 Matches: 648
 Percent Similarity: 97.44% Conservative: 0
 Best Local Similarity: 97.44% Mismatches: 2
 Query Match: 97.69% Indels: 15
 DB: 13 Gaps: 1
 US-09-581-651D-1 (1-660) x ADR67200 (1-8027)
 Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
 Db 214 AACTTGGTGGCAACTTGGCTCCCGGTGGGGCGTCTCTCCCCACCGCTCTCAACATGCTT 273
 Qy 21 ArgGlyProGlyProGlyLeuLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
 Db 274 AGGGGTCCGGGGCCCGGGTGGCTGGCTGGCGCGTCCAGTGGCGGACGCGGTGCC 333
 Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro 60
 Db 334 TCCACGGGAGCCTCGAAGACCAAGAGCAGGCTCAGCAATATGGTTACGCCCCCGCCG 393
 Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnIleAsn 80
 Db 394 GTGGCTGTCACTCAAGCAAGCCCGGTTGTATGACAAATGGAACACACTATCAGATNAAT 453
 Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 100
 Db 454 CAACAGTGGGAGCGGACCTTACCTAGGCAATAGCGTTGGTTGTACTTGTATGGAGGAAGC 513
 Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysTyrThr 120
 Db 514 CGAGGTTTTAACTCCGAGAGTAACCTGAAGACTGAGAGACTTGTCTTTGACAGTACACT 573
 Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140

Db 574 GGGAAACCTTACCGAGTGGGTGACACCTATGAGCGTCTTAAAGACTCCATGATCTGGGAC 633
Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAenArgCysHis 160
Db 634 TGTACCTGCATCGGGCTGGGCGAGGGAGGAATAAGCTGTACCATCGCAAAACCGCTGCCAT 693
Qy 161 GluGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180
Db 694 GAAGGGGGTTCAGTCTCAAGATTGGTGACACCTGGAGGAGACACATGAGACTGGTGGT 753
Qy 181 TyrMetLeuGluCysValCysLeuGlyAenGlyLysGlyGluTrpThrCysLysProIle 200
Db 754 TACATGTTAGAGTGTGTGTCTTGGTAATGGAAAAGAGNAATGGACCTCGAAGCCCATTA 813
Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
Db 814 GCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCTTATGTGGTGGGAGAACCTGGGAG 873
Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db 874 AAGCCCTACCAAGCTGGATGATGATGATTGTATCTTGCCTGGGAGAGGCGGACGC 933
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
Db 934 ATCACTTGCACTTCTAGAAATAGATGCAAGTCAAGACACACAGACATCTCTATAGAAAT 993
Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 994 GGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGATCTGCACAGGC 1053
Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1054 AACGCCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCGACACACATCAGCGGA 1113
Qy 301 SerGlyProPheThrAspValArgAlaValTyrGlnProGlnProHisProGlnPro 320
Db 1114 TCTGCCCCCTTCCCGATGTTCTGTGCAGCTGTTTACCAACCCGACGCTCACCCCCAGCCT 1173
Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
Db 1174 CCTCCTATGCCACTGTGTGCACAGACAGTGGTGTGTCTACTCTGTGGGATGCACTGG 1233
Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1234 CTGAAGACACAAAGGAATAAAGCAATGCTTTGACGTCCTGGGCAACGGAGTCAGCTGC 1293
Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCysValLeu 380
Db 1294 CAAGAGACAGCTGTMACCAGACTTACCGTGGCACTCAATGGAGACCATGTGCTTA 1353
Qy 381 ProPheThrTyrAsnAspArgThr----- 388
Db 1354 CCATTCACCTACATATGGCAGGACGTTCTACTCTGCACACGGAAGGGCGACGACGGA 1413
Qy 389 ----- AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
Db 1414 CATCTTTGGTGACGACCACTTCGAATATATGAGCAGGACCCAGAAATACTCTTTCTGCACA 1473
Qy 406 AspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCysHisPhe 425
Db 1474 GACCACTGTTTGGTTCAGACTCGAGGAGGAATTCNAATGGTGCCTTGTGCCACTTC 1533
Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgArgAspAsn 445
Db 1534 CCCTTCTCTATACAACACCAATTTACACTGATTCACCTTCTGAGGGCAGAGAGACAAAC 1593
Qy 446 MetLysThrCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
Db 1594 ATGAAGTGGTGGGACCAACACAGAACTATGATGCGGACCAAGATTGGGTTCGCCCC 1653
Qy 466 MetAlaAlaHisGluGluCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485

Db 1654 ATGGCTGCCACCAGGAAATCTGCACACCAATGAAGGGGTCTATGATCCGCATTGGAGAT 1713
Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
Db 1714 CAGTGGGATAAGCAGCATGACATGGGTGCATGATGAGGTGCACGTGTGTGGGAATGGT 1773
Qy 506 ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
Db 1774 CGTGGGGAATGGACATGCATTCCTACTCGACGTTCCGAGATCAGTGCATTGTTGATGAC 1833
Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545
Db 1834 ATCATTACAATCTGAACGACACATTCACACAGCGTCATGAAGAGGGGCGCATGCTGAAC 1893
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 565
Db 1894 TGTACATGCTTCGGTCAGGTCGGGCGAGTGGAAAGTGTATCCCTCGACCAATGCGCAG 1953
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrValHisGly 585
Db 1954 GATTCAGAGACTGGGACGTTTATCAATTTGGAGATTCATGGGAGAGTATGTGATGCT 2013
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu 605
Db 2014 GTCAGATACCAGTGTACTGCTATGCGCGTGGCAATTGGGAGTGGCAATGCCAACCTTTA 2073
Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2074 CAGACCTATCCAAAGCTCAAGTGTCTGTGCAAGTATTTATCACTGAGACTCCGAGTCAG 2133
Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2134 CCNACTCCCAACCCCATCCAGTGGATGCACACAGCCATCTCACATTTCCAAAGTACATT 2193
Qy 646 LeuArgTrpArgPro 650
Db 2194 CTCAGTGGAGACCT 2208
RESULT 13
ADQ38578
ID ADQ38578 standard; DNA; 6510 BP.
XX
XX AC ADQ38578;
XX
XX DT 18-NOV-2004 (first entry)
XX
XX DE Human SNP containing myocardial infarction-associated gene, SEQ ID 241.
XX
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX
XX KW cardiant; gene therapy; human; gene; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO2004058052-A2.
XX
XX PD 15-JUL-2004.
XX
XX PP 22-DEC-2003; 2003WO-US040978.
XX
XX PR 20-DEC-2002; 2002US-0434778P.
XX
XX PR 10-MAR-2003; 2003US-0453135P.
XX
XX PR 30-APR-2003; 2003US-0466412P.
XX
XX PR 23-SEP-2003; 2003US-0504955P.
XX
XX PA (APPL-) APPLERA CORP.
XX
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX
XX DR WPI, 2004-533949/51.
XX
XX DR P-PSDB; ADQ39406.
XX
XX PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in

Db 1819 CAGTGGGNTAAGCAGCATGACATGGGTCCACATGATGAGTGCACGTGTTGGGAATGGT 1878
 Qy 506 ArgGlyGluThrThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
 Db 1879 CGTGGGAATGGACATGATGCTTACTCGCAGCTTCGAGATCAGTGCATTTGTTGATGAC 1938
 Qy 526 IleThrTyrAanValAanAspThrPheHisIysArgHisGluGluGlyHisMetLeuAan 545
 Db 1939 ATCACTTACATGATGGAACGACACATTCACAAAGCGTCATGAAGGGGGGCACATGCTGAAC 1998
 Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 565
 Db 1999 TGTACATGCTTCGGTCAGGTCGGGGCAGGTGGAGTGTGATCCCGTCGACCAATGCCAG 2058
 Qy 566 AspSerGluThrGlyThrPheTyrClnIleGlyAspSerTrpGluLysTyrValHisGly 585
 Db 2059 GATTTCAGAGACTGGGACGCTTTTATCAATTCGAGATTCATGGGAGAGTATGTGCATGGT 2118
 Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluThrHisCysGlnProLeu 605
 Db 2119 GTTCAGATACCATGCTACTGTATGGCCGTGGCATTTGGGGAGTGGCATTTGCCAACCTTTA 2178
 Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
 Db 2179 CAGACCTATCCAGCTCAAGTGGTCTCTCGAGTATTTATCACTGAGACTCCGAGTCAG 2238
 Qy 626 ProAenSerHisProIleGlnTrpAanAlaProGlnProSerHisIleSerLysTyrIle 645
 Db 2239 CCCAACTCCCAACCCCATCCAGTGGGAATGCACACAGCCATCTCATATTTCCAAAGTACATT 2298
 Qy 646 LeuArgTrpArgPro 650
 Db 2299 CTCAGGTGGAGACCT 2313

RESULT 14
 ADQ38582
 ID ADQ38582 standard; DNA; 7823 BP.
 XX AC ADQ38582;
 XX 18-NOV-2004 (first entry)
 DT Human SNP containing myocardial infarction-associated gene, SEQ ID 245.
 DE Myocardial infarction; detection; single nucleotide polymorphism; SNP;
 KW cardiant; gene therapy; human; gene; ds.
 XX OS Homo sapiens.
 XX FN WO2004058052-A2.
 XX PD 15-JUL-2004.
 XX PF 22-DEC-2003; 2003WO-US040978.
 XX PR 20-DEC-2002; 2002US-0434778P.
 PR 10-MAR-2003; 2003US-0453133P.
 PR 30-APR-2003; 2003US-0466412P.
 PR 23-SEP-2003; 2003US-0504955P.
 XX (APPL-) APPLERA CORP.
 XX Cargill M, Devlin JJ, Iakubova O;
 DR WPI; 2004-533949/51.
 DR P-PSDB; ADQ39410.
 XX Identifying an individual who has an altered risk for developing
 PT myocardial infarction by detecting a single nucleotide polymorphism in
 PT the individual's nucleic acids.
 XX Claim 7; SEQ ID NO 245; 145pp; English.
 PS
 XX

CC The invention relates to a novel method for identifying an individual who
 CC has an altered risk for developing myocardial infarction. The method
 CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
 CC the nucleotide sequences given in the specification in the individual's
 CC nucleic acids, where the presence of the SNP is correlated with an
 CC altered risk for myocardial infarction in the individual. The invention
 CC further comprises: an isolated nucleic acid molecule comprising at least
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino
 CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting a SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiant activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC polynucleotide sequence represents a human myocardial infarction-
 CC associated gene containing one or more SNP's of the invention. Note: This
 CC sequence was not shown in the specification. The sequence has come from
 CC an electronic sequence listing downloaded from the WIPO website.
 XX
 SQ Sequence 7823 BP; 2066 A; 2031 C; 1935 G; 1776 T; 0 U; 15 Other;

Alignment Scores:
 Pred. No.: 3,06e-239 Length: 7823
 Score: 3687.50 Matches: 647
 Percent Similarity: 97.29% Conservatives: 0
 Best Local Similarity: 97.29% Mismatches: 3
 Query Match: 97.53% Indels: 15
 DB: Gaps: 1
 US-09-581-651D-1 (1-660) x ADQ38582 (1-7823)
 Qy 1 AenLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAanMetLeu 20
 Db 319 AACCTGGTGGCAACTTGCCTCCCGGTGGCGGTCTCTCCGCCACCGCTCTCAACATGCTT 378
 Qy 21 ArgGlyProGlyProGlyLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
 Db 379 AGGGGTCCGGGCGCCGGCTGCTGCTGCGCCGCTCCGCTGGGACAGCGGTGCC 438
 Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
 Db 439 TCCACGGGAGCCTCGAAGACGACAGAGGAGGCTCAGCAATGTTTCAGCCCGAGTCCCGC 498
 Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAenGlyLysHisTyrGlnIleAan 80
 Db 499 GTGGCTGTCACTCAAGCAAGCCCGTGTGTATGACAAATGGAAACACTATCAGATAAAT 558
 Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAanAlaLeuValCysThrCysTyrGlyGlySer 100
 Db 559 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTGTGTTGTACTTTGTTATGGAGGAAGC 618
 Qy 101 ArgGlyPheAenCysGluSerLysProGluAlaGluThrCysPheAspLysTyrThr 120
 Db 619 CGAGGTTTAACTCGGAGAGTAAACCTGAAGACTTTCGTTTGAACAAGTACACT 678
 Qy 121 GlyAenThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
 Db 679 GGGGAACACTTACCGAGTGGGTGACACTTATGAGCGTCTCTAAAGACTCCATGATCTGGGAC 738
 Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAanArgCysHis 160
 Db 739 TGTACCTGCATYGGGGCTGGCGGAGGAGAAATAGCTGTACCATCGCAACCGCTGCCAT 798
 Qy 161 GluGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180

Db 799 GAAGGGGTGAGTCTCTCAAGATTGGTGACACCTGGAGGAGACACATGAGACTGGTGGT 858
Qy TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluThrPheCysLysProile 200
Db TACATGTTAGAGTGTGTGTCTTGGTAATGGAAGAGGAGATGAGACTCGAAGCCCAT 918
Qy AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
Db CTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCTATATGTTGGTCGAGAAACCTGGGAG 978
Qy LysProTyrGlnGlyThrMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db AAGCCCTACCAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
Qy IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgile 260
Db ATCACTTGCACTTCTAGAAATAGATGCACATCAGGACACACAGACATCCTATAGAAAT 1098
Qy GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db GGAGACACCTGGAGCAAGAGATAATCGAGAAACCTGCTCCAGTGCATCTGCACAGGC 1158
Qy AsnGlyArgGlyGluThrLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db AACGGCCGAGAGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACACCATCGAGCGGA 1218
Qy SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
Db TCTGCCCCCTTACCGAGTTCGTGTCAGCTGTTTACCAACCGCAGCCTCACCCTCAGCCT 1278
Qy ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
Db CCTCCCTATGCCACTGTGTACAGACAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1338
Qy LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db CTGAAGACACAAAGAAATAAGCAATGCTTGTGCAGCTGCTGGGCAACGGAGTCAGCTGC 1398
Qy GlnGluThrAlaValThrGlnThrTyrGlyLysAsnSerAsnGlyGluProCysValLeu 380
Db CNAAGACAGCTGTATCCAGACCTTACGGTGGCACTCAATATGAGAGCCATGTGTCTTA 1458
Qy ProPheThrTyrAsnAspArgThr----- 388
Db CCATTCCACTACAATGGCAGGAGCTTCTACTCTCTGCACACAGAAAGGCGAGCAGCGGA 1518
Qy -----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
Db CATCTTTGGTGCAGCACAACACTTCGAATTTATGAGCAGGACCAAGAAATACTCTTTCTGCACA 1578
Qy AspHisThrValLeuValGlnThrArgGlyLysAsnSerAsnGlyAlaLeuCysHisPhe 425
Db GACCACACTGTTTGGTTCAGACTCGAGGAGAAATTCGAATGTTGCTGTGTGCGCACTTC 1638
Qy ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgArgAspAsn 445
Db CCTTCTCTATACACACCAACCAATTACATGATTGCACTTCTGAGGGCAGAGAGACAC 1698
Qy MetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
Db ATGAAGTGGTGTGGGACACACAGAACTATGATGCCAGCAGAAAGTTTGGGTTCTGCCCC 1758
Qy MetAlaAlaHisGluGluGluCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
Db ATGGCTGCCACAGGAGAAATCTGCACACCAACCAATGAAGGGGTCATGTACCGCATTTGGAGAT 1818
Qy GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
Db CAGTGGGATAAGCAGCATGATCGGTGCACATGATGAGTGCAGCTGTGTGTGGAAATGGT 1878
Qy ArgGlyGluThrTrpCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525

Db 1879 CGTGGGGAATGGACATGCATTGCTACTCGAGCTTCGAGATCAGTCATTGTTGATGAC 1938
Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545
Db 1939 ATCACTTTACAATGTGAACGACACATTCACAAGCGTCATGAAGAGGGGCACATGCTGAAC 1998
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 565
Db 1999 TGTACATGCTTCGGTCCAGGTCGGGCGCAGGTGGAAGTGTGATCCCGTCGACCAATGCCAG 2058
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrValHisGly 585
Db 2059 GATTCAAGACTGGGAGCTTTTATCAANTTGGAGATTCATGGGAGAGATATGTGCATGGT 2118
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu 605
Db 2119 GTCAGATACCAGTGTACTGCTATGCGCGTGGCATTTGGGAGTGGCATTTGCCAACCTTTA 2178
Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2179 CAGACCTTATCCAAGCTCAAGTGGTCTCTGTCGAAGTATTTATCACTGAGACTCCGAGTCAG 2238
Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2239 CCNACTCCCAACCCATCCAGTGGATGCCACAGCCATCTCACATTTCCAAAGTACATT 2298
Qy 646 LeuArgTrpArgPro 650
Db 2299 CTCAGTGGAGACCT 2313
RESULT 15
ADQ38583
ID ADQ38583 standard; DNA; 7848 BP.
XX AC ADQ38583;
XX AC
DT 18-NOV-2004 (first entry)
DE Human SNP containing myocardial infarction-associated gene, SEQ ID 246.
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; gene; db.
XX Homo sapiens.
XX OS
XX WO2004058052-A2.
XX PN
XX 15-JUL-2004.
XX PD
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX XX
PI Cargill M, Devlin JJ, Takoubova O;
XX WPI; 2004-533949/51.
DR DR
DR P-PSDB; ADQ39411.
XX XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX XX
PS Claim 7; SEQ ID NO 246; 145pb; English.
XX XX
CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's

CC nucleic acids, where the presence of the SNP is correlated with an
 CC altered risk for myocardial infarction in the individual. The invention
 CC further comprises: an isolated nucleic acid molecule comprising at least
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino
 CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiant activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC polynucleotide sequence represents a human myocardial infarction-
 CC associated gene containing one or more SNP's of the invention. Note: This
 CC sequence was not shown in the specification. The sequence has come from
 CC an electronic sequence listing downloaded from the WIPO website.
 xx
 SQ Sequence 7848 BP; 2069 A; 2046 C; 1939 G; 1778 T; 0 U; 16 Other;

Alignment Scores:

Pred. No.: 3,07e-239 Length: 7848
 Score: 3687.50 Matches: 647
 Percent Similarity: 97.29% Conservative: 0
 Best Local Similarity: 97.29% Mismatches: 3
 Query Match: 97.53% Indels: 15
 DB: 13 Gaps: 1

US-09-581-651D-1 (1-660) x ADQ38583 (1-7848)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
 Db 319 AACTTGGTGGCACTTGGCTCCCGGTGGGGGTCTCTCCCGCCGCTCTCAACATGCTT 378
 Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
 Db 379 AGGGTCCGGGGCCGGGCTGCTGCTGCTGGCCGTCWGTGCTGGGACAGCGGTGCC 438
 Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
 Db 439 TCCACGGGAGCTCGAAGAGCAAGAGGAGGAGCTCAGCAAAATGGTTCAGCCCGAGTCCCGC 498
 Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnIleAsn 80
 Db 499 GTGGCTGTCACTCAAGCAAGCCCGGTGTTATGCAATGGANAACATATCAGATAAT 558
 Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyCysSer 100
 Db 559 CAACAGTGGGAGCGGACCTACTAGGCAATGCGTTGGTTGTACTTTGTTATGGAGGAAGC 618
 Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysTyrThr 120
 Db 619 CGAGGTTTTAACTGCGAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACT 678
 Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
 Db 679 GGGAAACATTACCGAGTGGGTGACACTTATGAGCGTCTCTAAAGACTCCCATGATCTGGGAC 738
 Qy 141 CysThrCysIleGlyValaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
 Db 739 TGTACCTGCATVGGGGCTGGGCGAGGAGAGATAAGCTGTACCATCGCAAAACCGCTGCCAT 798
 Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyCly 180
 Db 799 GAAGGGGTGAGTCTCTACAAGATTGGTGACACCTCGAGGAGACCATGAGACTGGTGT 858
 Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysProIle 200

Db 859 TACATGTTAGAGTGTGTGTCTTGTGTAATGGAAGAGAGATGGACCTCAAGCCCAT 918
 Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
 Db 919 GCTGAGAAGTGTGTTGATCATGTCTGGGACTTCTTATGTGTGGAGAAACATGGGAG 978
 Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
 Db 979 AAGCCCTACCAAGGCTGGATGATGTTAGATTGTACTTGCCTGGGAGAGGACGCGGACG 1038
 Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
 Db 1039 ATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCTCTAGAAAT 1098
 Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
 Db 1099 GAGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC 1158
 Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
 Db 1159 ACGGCGAGAGAGTGGAGTGTGAGAGGCACACCTCTCTGCGAGACCACATCGAGCGGA 1218
 Qy 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
 Db 1219 TCTGGCCCTTCCACGATGTTCTGTCAGCTGTTTACCACCGCAGCCTCACCCCGACCT 1278
 Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
 Db 1279 CCTCCCTTATGGCAGCTGTGTCAGACAGAGTGGTGTGTCTACTCTGTGGGATGCATGG 1338
 Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
 Db 1339 CTGAGACACACAGGAATAAGCAATGCTTTGACGTCCTGGGCAACGGAGTCAGCTGC 1398
 Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCysValLeu 380
 Db 1399 CAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGAGAGACCATGTGTCTTA 1458
 Qy 381 ProPheThrTyrAsnAspArgThr----- 388
 Db 1459 CCATTTCACCTACCAATGGCAGGACGTTCTACTCTCTGCCACACAGAGGGCGACAGCGGA 1518
 Qy 389 -----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
 Db 1519 CATCTTTGGTGCACACCAACTTCGAATATGAGCAGACACAGAAATACTCTTTCTGCACA 1578
 Qy 406 AspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCysHisPhe 425
 Db 1579 GACCACACTGTTTTGGTTTCAGACTCGAGGAGGAAATCCAAATGGTGTCTTGTGCCACTTC 1638
 Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgAspAsn 445
 Db 1639 CCTTCTCTATACAAACCAACCAATTACACTGATTGCACTTCTGAGGGCAGAGAGACAAAC 1698
 Qy 446 MetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyCysPro 465
 Db 1699 ATGAAGTGGTGGGACCAACAGAACTATGATGCCACCAAGAGTTGGGTTCTGCCCC 1758
 Qy 466 MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
 Db 1759 ATGGCTCCCGCAGGAGAAATCTCGCAACCAATCAAGGGGTGATGTACCGCATTTGGAGAT 1818
 Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
 Db 1819 CAGTGGGATAAGCAGCATGATGGGTCAATGATGAGGTGCACGTGTGTGGGAAATGGT 1878
 Qy 506 ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
 Db 1879 CGTGGGATGGACATGCAATTCCTACTTCGAGAGTTTCGAGATCAGTGCATTTGTTGATGAC 1938
 Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2005, 22:35:24 ; Search time 367 Seconds
(without alignments)

9572.442 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 2147

Sequence: 1 caaactgtgtggcaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1897.4	87.9	8044	4	US-09-566-921-135
2	1826.8	85.1	7803	2	US-08-551-356-1
3	1826.8	85.1	7803	5	PCT-US93-12687-1
4	1743.4	81.2	7679	4	US-09-220-132-38
5	1743.4	81.2	7680	4	US-09-023-655-1289
6	1743.4	81.2	7680	5	PCT-US95-09819-6
7	1740.2	81.1	7705	1	US-08-259-569-16
8	1740.2	81.1	7705	2	US-08-826-885-16
9	1738.6	81.0	7705	6	5455158-2
10	1738.6	81.0	7705	6	5455158-2
11	671.2	31.3	986	1	US-07-637-250A-8
12	671.2	31.3	986	1	US-08-145-061-8
13	97.4	4.5	186	1	US-08-153-799-5
14	73.8	3.4	2109	4	US-09-799-451-345
15	73.8	3.4	2334	4	US-09-023-655-996
16	73.8	3.4	2334	4	US-09-949-016-704
17	73.8	3.4	2335	4	US-09-799-451-346
18	73.8	3.4	2335	4	US-09-949-016-4758
19	73.2	3.4	11665	4	US-09-949-016-12446
20	73.2	3.4	11665	4	US-09-949-016-16500
21	72.4	3.4	85	1	US-08-259-569-28
22	72.4	3.4	85	2	US-08-826-885-28
23	72.2	3.4	2334	1	US-08-457-304A-33
24	72.2	3.4	2334	1	US-08-456-701A-33
25	72.2	3.4	2334	3	US-08-684-932A-33
26	72	3.4	72	1	US-08-259-569-29
27	72	3.4	72	2	US-08-826-885-29

28	72	3.4	72	6	5455158-9	Patent No. 5455158
29	72	3.4	72	6	5455158-9	Patent No. 5455158
30	71.8	3.3	1983	4	US-09-949-016-2066	Sequence 2066, Ap
31	71.8	3.3	3069	4	US-09-949-016-641	Sequence 641, Appl
32	69	3.2	69	1	US-08-259-569-23	Sequence 23, Appl
33	69	3.2	69	2	US-08-826-885-23	Sequence 23, Appl
34	65.8	3.1	69	6	5455158-6	Patent No. 5455158
35	65.8	3.1	69	6	5455158-6	Patent No. 5455158
36	65.4	3.0	69	1	US-08-259-569-30	Sequence 30, Appl
37	65.4	3.0	69	2	US-08-826-885-30	Sequence 30, Appl
38	65.2	3.0	601	4	US-09-949-016-30391	Sequence 30391, A
39	65.2	3.0	601	4	US-09-949-016-70165	Sequence 70165, A
40	65.2	3.0	29954	4	US-09-949-016-13808	Sequence 13808, A
41	65.2	3.0	31040	4	US-09-949-016-12383	Sequence 12383, A
42	65	3.0	75	1	US-08-259-569-18	Sequence 18, Appl
43	65	3.0	75	2	US-08-826-885-18	Sequence 18, Appl
44	64.8	3.0	68	6	5455158-4	Patent No. 5455158
45	64.8	3.0	68	6	5455158-4	Patent No. 5455158

ALIGNMENTS

RESULT 1
US-09-566-921-135
; Sequence 135, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne P.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 135
; LENGTH: 8044
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 427813.14
US-09-566-921-135

Query Match	87.9%	Score 1887.4;	DB 4;	Length 8044;
Best Local Similarity	97.4%	Pred. No. 0;		
Matches 1946;	Conservative	0;	Mismatches	6;
Indels	45;	Gaps	1;	
Qy	1	CAAACTTGGTGGCAACTTGGCTCCCGTGGCGGCGTCTCTCCCGCACCGTCTCAACATGC	60	
Db	213	CAAACTTGGTGGCAACTTGGCTCCCGTGGCGGCGTCTCTCCCGCACCGTCTCAACATGC	272	
Qy	61	TTAGGGTTCGGGGCCGGGCTGCTGCTGGCGGCTCCAGTGTGGGACAGCGGTGC	120	
Db	273	TTAGGGTTCGGGGCCGGGCTGCTGCTGGCGGCTCCAGTGTGGGACAGCGGTGC	332	
Qy	121	CTTCCACGGGAGGCTCGAAGAGCAAGGCGAGCTCAGCAATGGTTTACGCCCGTCCC	180	
Db	333	CTTCCACGGGAGGCTCGAAGAGCAAGGCGAGCTCAGCAATGGTTTACGCCCGTCCC	392	
Qy	181	CGGTGGTGTGCTAGTCAAAAGCAAGCCGGTGTATGACAAATGGAACACATATCAGATAA	240	
Db	393	CGGTGGTGTGCTAGTCAAAAGCAAGCCGGTGTATGACAAATGGAACACATATCAGATAA	452	
Qy	241	ATCAACAGTGGGACGAGCCTAAGGCAATCGTGGTGTGTATCTTGTATGGAGGAA	300	
Db	453	ATCAACAGTGGGACGAGCCTAAGGCAATCGTGGTGTGTATCTTGTATGGAGGAA	512	
Qy	301	GCCGAGGTTTAACTCGGAGTAAACCTGAAGCTGAAGAGCTTGTCTTTGACAAGTACA	360	
Db	513	GCCGAGGTTTAACTCGGAGTAAACCTGAAGCTGAAGAGCTTGTCTTTGACAAGTACA	572	

QY 361 CTGGGAACACTTACCGAGTGGGTGACACTTATGACGCTCCTAAAGACTCCATGATCTGGG 420
DB 573 CTGGGAACACTTACCGAGTGGGTGACACTTATGACGCTCCTAAAGACTCCATGATCTGGG 632
QY 421 ACTGTACCTGATCGGGCTGGCGAGGGAGAAATAGCTGTACCACTGCAACCGCTGCC 480
DB 633 ACTGTACCTGATCGGGCTGGCGAGGGAGAAATAGCTGTACCACTCGCAACCGCTGCC 692
QY 481 ATGAAGGGGTGAGTCTTACAAGATTGGTGACACTGGAGAGAGACCAATGAGACTGGTG 540
DB 693 ATGAAGGGGTGAGTCTTACAAGATTGGTGACACTGGAGAGAGACCAATGAGACTGGTG 752
QY 541 GTTACATGTTAGAGTGTGTCTTGTGTTATGGAAGAGGAAATGGACCTGCAAGCCCA 600
DB 753 GTTACATGTTAGAGTGTGTCTTGTGTTATGGAAGAGGAAATGGACCTGCAAGCCCA 812
QY 601 TAGCTGAGAAGTGTGTTGATCATGCTGTGGGACTTCTTATGTGTCGGAGAAACGTGGG 660
DB 813 TAGCTGAGAAGTGTGTTGATCATGCTGTGGGACTTCTTATGTGTCGGAGAAACGTGGG 872
QY 661 AGAAGCCCTACCAAGGCTGATGATGGTAGATTGTTGCTGGGAGAAAGCGAGCGAC 720
DB 873 AGAAGCCCTACCAAGGCTGATGATGGTAGATTGTTGCTGGGAGAAAGCGAGCGAC 932
QY 721 GCATCATTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTTATAGAA 780
DB 933 GCATCATTGCATCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTTATAGAA 992
QY 781 TTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAG 840
DB 993 TTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAG 1052
QY 841 GCAACGGCGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGTCAGACCAATCGAGCG 900
DB 1053 GCAACGGCGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGTCAGACCAATCGAGCG 1112
QY 901 GATCTGGCCCTTACACCGATGTTGTCGAGCTGTTTACCAACCGCAGCTCACCCCCAGC 960
DB 1113 GATCTGGCCCTTACACCGATGTTGTCGAGCTGTTTACCAACCGCAGCTCACCCCCAGC 1172
QY 961 CTCCTCCCTATGGCCACTGTGTACAGACAGTGGTGTCTACTCTCTGTGGGATGCAGT 1020
DB 1173 CTCCTCCCTATGGCCACTGTGTACAGACAGTGGTGTCTACTCTCTGTGGGATGCAGT 1232
QY 1021 GGCTGAAGACACAGGAAATAGCAAAATGCTTTGACAGCTGCTGGGCAACGGAGTCAGT 1080
DB 1233 GGCTGAAGACACAGGAAATAGCAAAATGCTTTGACAGCTGCTGGGCAACGGAGTCAGT 1292
QY 1081 GCCAAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAATGGAGGCCATGTGTCT 1140
DB 1293 GCCAAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAATGGAGGCCATGTGTCT 1352
QY 1141 TACCATTACCTTACCAACGACAGGAC----- 1165
DB 1353 TACCATTACCTTACCAATGGCAGGACGTTCTACTCTGCACACACAGAGGGCGACAGGACG 1412
QY 1166 -----GGACAGCAACCTTCGAATTTATGACAGGACAGGAAATACTCTTTCTGCA 1215
DB 1413 GACATCTTTGGTGACAGCAACCTTCGAATTTATGACAGGACAGGAAATACTCTTTCTGCA 1472
QY 1216 CAGACCACTGTTTGGTTGAGCTCGAGGAGGAAATCCAAATGGTGCCTTGTGCCACT 1275
DB 1473 CAGACCACTGTTTGGTTGAGCTCGAGGAGGAAATCCAAATGGTGCCTTGTGCCACT 1532
QY 1276 TCCCTCTCTTACCAACCAACCAATTTACCTGATTGCACTTCTGAGGGCGAGAGAGACA 1335
DB 1533 TCCCTCTCTTACCAACCAACCAATTTACCTGATTGCACTTCTGAGGGCGAGAGAGACA 1592
QY 1336 ACATGAAGTGGTGGGACCAACAGAACTATGATGCCGACAGCAAGTTTGGGTTCTGCC 1395
DB 1593 ACATGAAGTGGTGGGACCAACAGAACTATGATGCCGACAGCAAGTTTGGGTTCTGCC 1652

QY 1396 CCATGGCTGCCACGAGGAAATCTGCAACCAATGAAGGGGTCACTGTACCGCATTTGGAG 1455
DB 1653 CCATGGCTGCCACGAGGAAATCTGCAACCAATGAAGGGGTCACTGTACCGCATTTGGAG 1712
QY 1456 ATCAAGTGGGTAAGCAGCAGTGAACATGGGTCAATGATGAGGTGCACTGTGTTGGGAATG 1515
DB 1713 ATCAAGTGGGTAAGCAGCAGTGAACATGGGTCAATGATGAGGTGCACTGTGTTGGGAATG 1772
QY 1516 GTCGTGGGTAAGCAGCAGTGAACATGGGTCACTGCGAGCTTCGAGATCACTGATTTGCATG 1575
DB 1773 GTCGTGGGTAAGCAGCAGTGAACATGGGTCACTGCGAGCTTCGAGATCACTGATTTGCATG 1832
QY 1576 ACATCACTTAAATGTGAACGACACACATTTCCACAAGCGCTCATGAAGAGGGGACATGCTGA 1635
DB 1833 ACATCACTTAAATGTGAACGACACACATTTCCACAAGCGCTCATGAAGAGGGGACATGCTGA 1892
QY 1636 ACTGTAATGCTTCGGTTCAGGGTCGGGGCAGGTGGAAAGTGTGATCCCGTCGACCAATGCC 1695
DB 1893 ACTGTAATGCTTCGGTTCAGGGTCGGGGCAGGTGGAAAGTGTGATCCCGTCGACCAATGCC 1952
QY 1696 AGGATTCAGAGACTGGGACGTTTATCAAAATGGAGATTCATGGGAGAAATGATGTCATG 1755
DB 1953 AGGATTCAGAGACTGGGACGTTTATCAAAATGGAGATTCATGGGAGAAATGATGTCATG 2012
QY 1756 GTGTCAATACCAAGTGTCTACTGCTATGCGCGTGGCATTTGGGAGTGGCAATTCGCAACCTT 1815
DB 2013 GTGTCAATACCAAGTGTCTACTGCTATGCGCGTGGCATTTGGGAGTGGCAATTCGCAACCTT 2072
QY 1816 TACAGACTTATCAAGCTCAAGTGGTCTGTGCAAGTATTTATCACTGAGACTCCGAGTC 1875
DB 2073 TACAGACTTATCAAGCTCAAGTGGTCTGTGCAAGTATTTATCACTGAGACTCCGAGTC 2132
QY 1876 AGCCCAACTCCCAACCCCATCCAGTGGAAATGCACACAGCCATCTCAGATTTCCAAAGTACA 1935
DB 2133 AGCCCAACTCCCAACCCCATCCAGTGGAAATGCACACAGCCATCTCAGATTTCCAAAGTACA 2192
QY 1936 TTCTCAGTGGAGACCT 1952
DB 2193 TTCTCAGTGGAGACCT 2209

RESULT 2

US-08-551-356-1
; Sequence 1, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Holly, Julie A
; REGISTRATION NUMBER: 33-246

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 17670.7 Seconds
(without alignments)
1809.796 Million cell updates/sec

Title: US-09-581-651D-1

Perfect score: 3781

Sequence: 1 NLVATCLFVRASLPHRLNML.....ISKYILRWPRVSPRPNLGY 660

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092223_28778/app_query.fasta_1.1834
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -WATRIK=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651@cgn 1.1 9936 @runat_07112005_092223_28778 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3781	100.0	2147	6 BD137021	BD137021 Polypepti
2	3781	100.0	2147	6 AX003229	AX003229 Sequence
3	3781	100.0	2147	9 HSA276395	AU276395 Homo sapi
4	3774	99.8	2358	6 CQ875358	CQ875358 Sequence

ALIGNMENTS

RESULT 1	BD137021	2147 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	BD137021				
DEFINITION	Polypeptides, polynucleotides and uses thereof.				
ACCESSION	BD137021				
VERSION	BD137021.1	GI:23231966			
KEYWORDS	JP 2002508179-A/1.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Schor, S.L. and Schor, A.M.				
TITLE	Polypeptides, polynucleotides and uses thereof				
JOURNAL	Patent: JP 2002508179-A 1 19-MAR-2002;				
COMMENT	UNIVERSITY OF DUNDEE OS Homo sapiens (human) PN JP 2002508179-A/1 PD 19-MAR-2002 PF 15-DEC-1998 JP 2000539133 PR 16-DEC-1997 GB 9726539.1 PI SETH LAWRENCE SCHOR, ANA MARIA SCHOR PC C12N15/09, A61K38/00, A61P17/02, C07K14/78, C07K16/18, C12N5/10, PC C12P21/08, PC C12Q1/68, G01N33/574, C12N15/00, A61K37/02, C12N5/00 CC Polypeptides, polynucleotides and uses thereof. PH Key Location/Qualifiers 1. .2147 FT source				

AJ535086 Homo sapi
BX538045 Homo sapi
AB191261 Homo sapi
CQ833991 Sequence
CQ875357 Sequence
AL833771 Homo sapi
BX640608 Homo sapi
BX640802 Homo sapi
BX641182 Homo sapi
BX454662 Sequence
AX281712 Sequence
BX640875 Homo sapi
BX538017 Homo sapi
BX538018 Homo sapi
BX640731 Homo sapi
CQ871810 Sequence
CQ871828 Sequence
AL832202 Homo sapi
BX537590 Homo sapi
BX640803 Homo sapi
AR051657 Sequence
AR274901 Sequence
AR380744 Sequence
AX227596 Sequence
AX333368 Sequence
X02761 Human mRNA
A14133 Fibronectin
E01162 cDNA encodi
E01162 Sequence 16
I70110 Sequence
AK364992 Sequence
BC051082 Mus muscu
AX402055 Sequence
X15906 Rat mRNA fo
CQ715726 Sequence
CQ72841 Xenopus l
M77820 Xenopus lae
BV178397 sqm10026
AF081127 Danio rer
AF081128 Danio rer
AY538257 Danio rer

FEATURES	FT	Location/Qualifiers	/organism='Homo sapiens (human)'
Source		1..2147	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Alignment Scores:			
Pred. No.:	1.1e-263	Length:	2147
Score:	3781.00	Matches:	660
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-581-651D-1 (1-660) x BD137021 (1-2147)			
Qy	1	AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu	20
Db	3	AAC TTGGTGGA CACTTGCTCCCGGTGGGGGGTCTCTCCCCACCGTCTCAACATGCTT	62
Qy	21	ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro	40
Db	63	AGGGGTCCGGGCGCGGGCTGTCTGTCTGGCCGCTCCAGTGCCTGGGACAGCGGTGCC	122
Qy	41	SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro	60
Db	123	TCCACGGGAGGCTCGAAGACAAGAGCAGGCTCAGCAAAATGGTTCAGCCCCAGTCCCGC	182
Qy	61	ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnLeu	80
Db	183	GTGGCTGTCAAGTCAAAAGCAAGCCCGGTGTTATGACAATGGAAAAACATATCAGATAAAT	242
Qy	81	GlnGlnThrGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer	100
Db	243	CAACAGTGGGACGGACCTACTAGCGCAATGCGTTGGTTGTACTGTTATGGAGGAAGC	302
Qy	101	ArgGlyPheAsnCysGluSerLysProGluAlaGluThrCysPheAspLysTyrThr	120
Db	303	CGAGGTTTAACTGGAGAGTAAACCTGAAGCTGAAGAGACTTGCTTTGACAAGTACACT	362
Qy	121	GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp	140
Db	363	GGGAACACTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATGTGGGAC	422
Qy	141	CysThrCysIleGlyValaclyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis	160
Db	423	TGTACCTGCATCGGGGCTGGGCGAGGAGAAATAGCTGTACCATTGCCAAACCGCTGCCAT	482
Qy	161	GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly	180
Db	483	GAAGGGGTTCAGTCTCTACAAGATTGGTGACACTGGAGGAGACCATGAGACTGGTGTGT	542
Qy	181	TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyValTrpThrCysLysProIle	200
Db	543	TACATGTTAGAGTGTGTGTCTCGTAAATGGAAAAGGAGAAATGGACCTGCAAGCCCAT	602
Qy	201	AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu	220
Db	603	GCTGAGAAGTGTTTTGATCATGCTGCTGGACTTCTCTATGTGGTGGAGAAACGTGGGAG	662
Qy	221	LysProTyrGlnGlyTyrMetMetValAspCysThrCysLeuGlyValGlySerGlyArg	240
Db	663	AAGCCCTTCAAGGCTGGATGATGGTAGATTCTGTACTTGCCTGGGAGAAAGCGACGCGC	722
Qy	241	IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle	260
Db	723	ATCACTTGACATTCTAGAAATAGATCAACATCAGACACAAAGGCATCTCTATAGAAAT	782
Qy	261	GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly	280
Db	783	GGAGACACTGGAGCAAGAGGATTAATCAGGAAACCTCTCCAGTGCATCTGCACAGCG	842